PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 17 June 15, 1931 Number 6

IMMUNITY TO SUGAR CANE MOSAIC ACQUIRED BY THE HOST

By E. M. East

Harvard University, Bussey Institution

Communicated May 1, 1931

Some ten or twelve years ago Mr. R. M. Grey, of the Harvard Botanic Gardens in Cuba, made a series of observations which convinced him that sugar cane infected with the virus of the mosaic disease can throw off all pathological symptoms and remain apparently healthy for a time. During the month of February, 1924, Dr. W. H. Weston, Jr., and I were able to verify many of these observations.

Some months before this date, Mr. Grey planted an isolated plot of infected canes in the Botanic Garden and gave them the best cultural treatment that he could devise, hoping that the plants would show some evidence of recovery during our visit. This hope was realized. Canes were found in all stages of what appeared to be a true convalescence. The older, dead and dying leaves exhibited the characteristic markings of mosaic infection, while the younger leaves, in the vicinity of the growing point, were apparently healthy. This appearance contrasted markedly with that of plants in other fields which were taking on the symptoms of the disease. Here the younger leaves showed unmistakable evidence of infection, while the older leaves remained unmarked.

Evidence that the disappearance of all symptoms of mosaic infestation is not uncommon was found in fields which had shown between 70 per cent and 90 per cent of infected plants during 1923. In 1924 these same fields showed only between 40 per cent and 60 per cent of infected plants, though the loss of plants was negligible. The records of Mr. Grey contained many examples of such apparent recovery. In February, 1930, six years later, I had the privilege of examining the plants of the experimental plot for the second time. I found numerous plants which had been noted previously as having completely recovered from infection, if one accepts the ordinary external symptoms as truly diagnostic. How long these plants had remained in apparent health, I am unable to say; but, at this second observation date, they had the disease and were again recovering from it.

Mr. Grey, whom I have found to be a very accurate observer, informed me that he has noted as many as three successive infections and recoveries in a given plant. The periods during which the plants may exhibit disease symptoms and freedom from them (immunity?) are variable. There may be no recovery after an infection; or, there may be recovery at any time between three months and twenty-four months. After recovery, a plant may again show evidence of infection in somewhat less than a year; or, it may remain healthy for several years (no observation period having been longer than four years). Records of individual plants have not been sufficiently extensive to determine the extreme limits of time during which this apparent immunity persists.

As I suggested in an earlier paper (East and Weston, 1925),[1] such phenomena are interpretable by either of two hypotheses, both of which are immunological in character. The host may kill the infective agent, throw off all symptoms of the disease, remain in a partially immune condition for a period, and then become reinfected. It is equally probable that the host merely reduces the virulence of the mosaic virus until the latter is unable to produce the usual mosaic symptoms, although the virus continues to live within its tissues. In this case the visible symptoms of the disease could reappear without reinfection if the resistance of the plant were lowered by the proper combination of external and internal conditions.

From the theoretical point of view it is immaterial as to which of these hypotheses proves to be the more acceptable. The important fact is that here, for the first time, there seems to be critical evidence of the establishment of an acquired immunity to a definite infective agent in one of the flowering plants. But to the sugar-cane grower there is a great difference between the two assumptions. If cane plants can throw off the disease completely, a reinfection is essential for a second appearance of the symptoms. If, on the other hand, the reactions of the host plant simply reduce the effectiveness of the virus and yet permit it to remain and multiply, then there is no criterion by which to determine the prevalence of the disease. It is possible, indeed, that practically all of the individuals of a susceptible variety are mosaic carriers.

So far as I am aware, no observations have been made which permit one to make a critical decision. The existing evidence fits either hypothesis. Even if one accepts the idea that the transmission of sugar cane mosaic takes place solely by means of insect vectors—chiefly, or perhaps invariably, through *Aphis maydis*—as indicated by the laboratory work of Brandes and others and by the high incidence of the disease around the borders of a field, there is still no proof that the majority of apparently healthy plants are not carriers; for the insect vector may have the capacity for increasing the virulence of the mosaic culture.

In order to try to decide whether or not the proteins of various types

of canes were different, a series of precipitin experiments were carried out. Ten-gram samples were ground up in a mortar with twenty cubic centimeters of distilled water. The extract was filtered quickly under pressure at a temperature of about 10°C. (ice pack). The various extracts were tested against each other by the Uhlenhuth ring method.

Various canes, including young Cristalina non-mosaic, young Cristalina mosaic, old Cristalina non-mosaic, old Cristalina mosaic, Cristalina having mosaic for second time, Cristalina seedling non-mosaic, Cristalina seedling mosaic, and C 64 (a cane nearly immune to mosaic), were tested against each other in all possible combinations. All tests were negative. No precipitin rings were formed. It is to be noted, however, that all of these canes, with the possible exception of C 64, may have been carrying mosaic.

A second experiment consisted in testing *Saccharum biflorum* and *Saccharum ciliare*—two mosaic immune types which constantly reacted plus 1 against each other—against (*a*) badly infected Cristalina stock, (*b*) non-infected Cristalina stock which, nevertheless, may have been carrying mosaic, and (*c*) a Cristalina seedling known never to have had mosaic. No precipitin reactions were obtained against "*a*" and "*b*," but consistent plus 1 reactions were obtained against "*b*." Consistent plus 1 reactions were also obtained when "*a*" or "*b*" were tested against "*c*."

In a third experiment, the following canes of about the same age were used: "*d*," mosaic Cristalina, "*e*," non-mosaic Cristalina, and "*f*," a seedling H21087 known never to have had mosaic. These canes were tested individually against extracts of *Cucumis anguria* L., *Carica papaya* L., *Hibiscus archeri* Hort., *Lycopersicum esculentum* L., *Anona reticulata* L., *Phaseolus lunatus* L., *Eriobotrya japonica* Lindl., *Lucuma nervosa* A. DC., and *Cassia spectabilis* DC. With the first six species listed, there was no reaction, which perhaps makes somewhat questionable the specificity of the precipitin reaction. Tested with *Lucuma nervosa*, the reaction was plus 1 in each case. Tested with *Eriobotrya japonica* and *Cassia spectabilis*, the reaction was consistently plus 1 in the case of the cane which was known never to have had mosaic and consistently plus 2 in the cases of the mosaic Cristalina and of the non-mosaic Cristalina.

It is not my intention to claim any critical diagnostic importance for these precipitin tests. We do not know just what their significance is, either chemically or serologically. It seems legitimate to point out, however, that in these tests there is some indication of a consistent reaction difference between Cristalina type canes known never to have had the mosaic disease and Cristalina type canes which have or may have had the mosaic disease. In other words, plants which exhibit the mosaic disease react in the same manner as plants which appear to be free from the disease, though they may be acting as carriers. These reactions are

different from those shown by canes of similar stock which are known never to have had mosaic. Slight as the evidence undoubtedly is, it points toward the truth of the second of our hypotheses; that is to say, it is perhaps more probable that sugar-cane plants gain an apparent immunity by reducing the virulence of the mosaic virus than by throwing it off entirely.

[1] East, E. M., and Weston, W. H., Jr., "A Report on the Sugar Cane Mosaic Situation in February, 1924, at Soledad, Cuba," *Harv. Univ. Press*, Cambridge, 1925, pp. 52, plates IX.

CYTOLOGICAL EVIDENCE FOR SEGMENTAL INTERCHANGE BETWEEN NON-HOMOLOGOUS CHROMOSOMES IN MAIZE[1,2]

BY D. C. COOPER AND R. A. BRINK

DEPARTMENT OF GENETICS, UNIVERSITY OF WISCONSIN

Communicated May 11, 1931

During the past few years a number of cases of partial sterility in maize have come to light which appear to involve structural changes in the chromosomes. In one of these, semisterile-1, clear-cut genetic evidence has been secured indicating that the phenomenon is associated with alterations in two linkage groups of the nine which it is now possible to test.[3,4] Burnham[5] reports that at diakinesis in three different semisteriles, -1, -2 and -3, a ring of four chromosomes and eight bivalents occurs. The rings in semisteriles-1 and -3 appear to have one member in common, since crosses between the two lines give some offspring with a group of six chromosomes and seven bivalents. Semisterile-2 in combination with semisterile-1, on the other hand, gives two rings of four chromosomes each plus six bivalents.

Critical genetic evidence regarding the character of the change in the nuclear complement leading to ring formation at diakinesis is not yet forthcoming. It is anticipated that alterations in the composition of certain linkage groups corresponding to the particular chromosomes concerned will be found, and studies to this end are in progress. On the basis of observations on a similar phenomenon in other genera, Belling[6] has brought forward the hypothesis that ring-formation at diakinesis is a consequence of interchange of terminal segments between non-homologous chromosomes. Assuming attachment of homologous ends, a normal and an interchanged chromosome of each pair would give a closed group of four. As Burnham[5] has pointed out, however, a ring might be expected following a simple translocation if a cross-over occurred in the "four-strand" stage.

In a recent paper, Miss McClintock[7] has presented strong cytological evidence for segmental interchange in semisterile-2. The chromosomes affected are the second and third smallest. A study of the early prophase stages, in which the thread-like chromosomes are paired throughout their entire length, revealed a cross-shaped structure in the semisterile plants. One of the pairs of chromosomes in this complex, in the strain employed, possessed characteristic terminal knobs which serve to identify it and the interchange group. Utilizing these morphological features in the analysis, Miss McClintock has shown that the cross-shaped configuration is very probably the result of segmental interchange between the second and third smallest chromosomes.

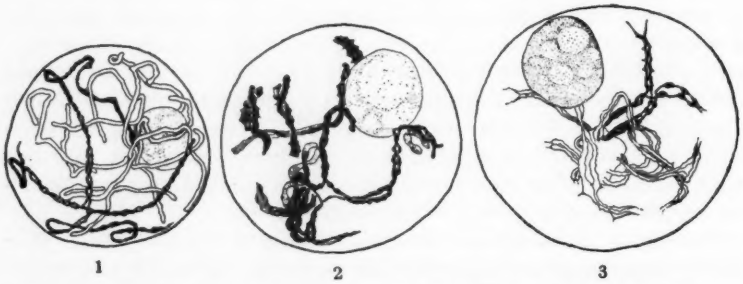
Our own investigations have been concerned with two other stocks, semisterile-1 and -5. Semisterile-5 is a particularly favorable object for this study in that one of the two pairs of chromosomes in the ring is regularly attached (or closely appressed) to the nucleole at the end which bears satellites.[8] By means of this feature this pair of chromosomes is readily distinguished from the other nine. Semisterile-5 in combination with semisterile-1 shows a ring of six chromosomes and seven bivalents. The two semisteriles, therefore, have one chromosome in common, that corresponding to either the *B-lg* or the *P-br* linkage group.

Portions of the staminate inflorescence from field-grown plants were fixed in a modification of Carnoy's solution which contained 6 parts 95 per cent alcohol, 2 parts glacial acetic acid and 2 parts chloroform. The material was allowed to remain in the fixing fluid for 15–30 minutes and was then transferred to 90 per cent alcohol. After a period of 12–24 hours it was stored in 80 per cent alcohol. The mounts were made in freshly prepared aceto-carmine to which had been added a fourth part of Ehrlich's haematoxylin. After crushing an anther in the stain a cover glass was added to the mount and the slide slightly warmed. The preparations, when sealed by running melted paraffin around the edge of the cover glass, will keep for a month without deterioration. Permanent slides were made in the manner described by Miss McClintock.[9]

On the basis of the segmental interchange hypothesis, the ring-forming group in semisterile individuals comprises two normal chromosomes (one of each pair) and two chromosomes which are complementary to each other as a result of having exchanged terminal segments. In the open spireme stage in normal plants, the members of a chromosome pair are closely associated throughout their whole length, homologous portions presumably lying in juxtaposition. At this stage the interchange complex would be expected to form a cross-shaped figure since each of the complementary chromosomes should synapse for a part of its length with one normal chromosome and for the remainder of its length with the other normal chromosome. On the basis of a simple translocation a linear configuration

or a T-shaped structure should occur, depending upon whether the detached segment was affixed in its new position by the broken or the free end.

Examination of a number of preparations in the open spireme stage and the stages immediately following reveals that semisterile-5 plants regularly show a cross-shaped chromosome complex attached by one end to the nucleole (Fig. 1). Several figures obtained were sufficiently clear to permit of analysis. One arm of the cross, bearing statellites, is affixed to the nucleole. The other three arms show a definite size relationship to each other, one being comparatively short, another of medium length and the third longer. The arm of medium length always lies opposite the attached arm.



1 2 3

Figure 1.—Nucleus of a microspore mother cell of semisterile-5 shortly after the open spireme stage. The cross-shaped figure with one arm attached to the nucleole is clearly apparent. ×775.

Figure 2.—Early diakinesis in semisterile-5 showing eight bivalents and the cross-shaped figure beginning to open into a ring. The arm opposite the attached one turns downward and therefore appears foreshortened in the figure. ×1450.

Figure 3.—Nucleus of microspore mother cell of semisterile-1 shortly after synizesis. The cross-shaped figure shows an exchange of partners. ×1450.

Six cross-shaped figures at stages from the open spireme to mid-diakinesis were measured. The lengths of the three arms, taking the one opposite the attached end as unity, were as follows:

0.5	1.0	1.3
0.5	1.0	1.4
0.6	1.0	1.6
0.7	1.0	1.5
0.4	1.0	1.9
0.4	1.0	1.4

At very early diakinesis eight bivalent chromosomes and the cross-shaped complex may be recognized, as shown in figure 2. In this figure the arm opposite the attached one turns downward at the end for a distance

of about 6μ and consequently appears foreshortened in the drawing. At this stage the regions of spindle fibre attachment appear as short hyaline areas in the four chromosomes which later form the ring. In normal plants the spindle fiber attachment on the satellite chromosome appears to be at a point about one-quarter of the length of the chromosome from the satellite end. In the interchange complex, therefore, it is the longer chromosome attached to the nucleole which is the normal satellite chromosome of the set; and the chromosome opposite this is the normal member of the other pair in the group. The interchange appears to have involved approximately two-thirds of the satellite chromosome and about one-third of the other chromosome entering into the ring. These relations are depicted diagrammatically in figure 4.

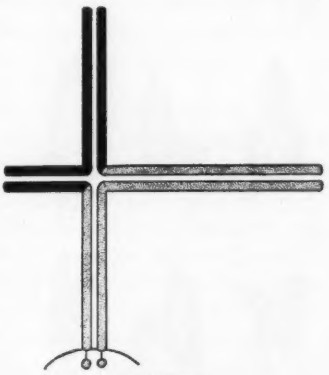


FIGURE 4

A diagrammatic representation of the cross-shaped figure found in semisterile-5. Taking the length of the arm opposite the satellites as unity, the lengths of the two laterals are approximately 0.5 and 1.5.

At the first meiotic division three types of distribution of the chromosomes in the semisterile-5 ring are observed. Adjacent chromosomes of approximately the same size may go to the same pole (Fig. 5a), adjacent chromosomes of different sizes may assort together (Fig. 5b) or alternate chromosomes may pass to the same pole (Fig. 5c). Only the spores resulting from the last type of distribution will be functional since the others lack at least a portion of one chromosome. Counts show about 50 per cent aborted pollen in semisterile-5 plants. This means that in approximately half the cases alternate chromosomes in the ring assort together.

While semisterile-1 is a less favorable object for cytological study on account of the absence of easily distinguishable morphological features in the chromosomes concerned, some evidence has been obtained that its behavior is also due to segmental interchange between non-homologous chromosomes. By means of breeding tests it has been possible to show in this case that the chromosomes affected correspond to the *B-lg* and *P-br* linkage groups. Burnham[5] has reported the occurrence of a chromosome ring plus eight bivalents in this stock, an observation amply confirmed by our own studies.

In early diakinesis figures the interchange complex in semisterile-1 is easily recognized inasmuch as the group of four chromosomes comes to lie at one side of the nucleus and apart from the remaining bivalents. Four

arms consisting of paired strands can be seen. One of the arms is composed of a row of three chromatic bodies. Since neither pair of chromosomes concerned in semisterile-1 is easily identified, it is difficult to trace the cross-shaped structure in the early prophase stages. In figure 3 is shown, however, a very early spireme stage in which a configuration occurs suggesting segmental interchange. The cross-shaped structure is made up of four arms, each arm being composed of paired strands. At the midpoint of this synaptic complex two of the strands exchange partners. Here, as in several early diakinesis figures of semisterile-1 which we have observed, one of the arms consists of a row of three paired chromatic bodies. It seems probable that this configuration is the precursor of the ring seen at diakinesis, but the evidence is less decisive than that bearing on semisterile-5.

FIGURE 5

Chromosome rings from heterotypic equatorial plates of semisterile-5 showing the distribution of the component chromosomes: a, adjacent chromosomes of unequal size going to the same pole; b, adjacent chromosomes of about the same size going to the same pole; c, alternate chromosomes going to the same pole. ×1450.

[1] Papers from the Department of Genetics, Agricultural Experiment Station, University of Wisconsin, No. 122. Published with the approval of the Director of the Station.

[2] The authors desire to express their appreciation for the assistance received from the Research Fund of the University of Wisconsin, and through a grant-in-aid from the NATIONAL RESEARCH COUNCIL.

[3] Brink, R. A., *Anat. Record* (abstract), **44**, 280 (1929).

[4] Brink, R. A., and Cooper, D. C., *Genetics* (in press).

[5] Burnham, C. R., *Proc. Nat. Acad. Sci.*, **16**, 269–277 (1930).

[6] Belling, J., *Zeit. indukt. Abstamm. Vererb.*, **39**, 286–288 (1925).

[7] McClintock, B., *Proc. Nat. Acad. Sci.*, **16**, 791–796 (1930).

[8] McClintock, B., *Science*, **69**, 629 (1929).

[9] McClintock, B., *Stain Tech.*, **4**, 53–56 (1929).

THE BEHAVIOR OF THE VIRUS OF YELLOW FEVER IN MONKEYS AND MICE

BY ANDREW WATSON SELLARDS*

DEPARTMENT OF TROPICAL MEDICINE, HARVARD MEDICAL SCHOOL

Communicated May 15, 1931

The earlier efforts to find a laboratory animal that would react in any way to inoculation with yellow fever resulted only in repeated failures. Now it appears, rather unexpectedly, that several distinct species are readily susceptible. Stokes, Bauer and Hudson[1] opened an important field of work by the discovery that the rhesus monkey is a very satisfactory animal for the study of this disease. One of the most interesting observations in this new field of work was made by Theiler[2] who showed that white mice die regularly after an intracerebral inoculation of the virus of yellow fever. Mice injected in this manner develop a fatal encephalitis but the liver remains normal whereas the disease in man and monkeys is characterized by extensive necrosis of the liver. Furthermore, Theiler noted that the virus, after repeated passages in mice, loses its virulence for monkeys upon injection in the usual manner into the peritoneum. This loss of infectivity for the monkey obviously makes it necessary to determine whether the strain of yellow fever which was originally inoculated into the mice may have become contaminated or supplanted by some unknown virus existing in a latent condition in stock mice or monkeys. Theiler provided evidence that the encephalitis in mice is due to yellow fever by showing that the serum of patients convalescent from yellow fever affords a measurable degree of specific protection. Subsequently, the observations of Theiler have received full confirmation notably in the laboratories of the International Health Board.

The question of the identity of the disease, producing encephalitis in mice without necrosis of the liver is of fundamental importance in any study of the etiology of yellow fever. Direct proof that the infection in mice is yellow fever was sought in the following experiments by attempting to restore its virulence for monkeys. Also, a series of cross immunity tests were carried out. A typical strain of yellow fever maintained in monkeys and mosquitoes and producing necrosis of the liver was compared with the "mouse virus" which is this same strain of virus modified by repeated passage from brain to brain of mice unless it has inadvertently been supplanted by some unknown virus.

Intracerebral Injection of Monkeys.—A suspension of infective mouse brain which was avirulent for monkeys upon subcutaneous or intraperitoneal inoculation, was injected into the left anterior frontal lobe of the brain of a monkey (*Macacus rhesus*). On the sixth day after injection,

fever developed (104.5°F.) accompanied by an almost complete paralysis of the right side of the body suggesting focal growth of the virus in the brain. Death occurred on the seventh day. The gross anatomical findings were not remarkable. The subcutaneous tissues were distinctly yellow, the liver was red and firm, the stomach and spleen were normal in appearance. Twenty cc. of blood were withdrawn from the heart of this animal just before death and injected intraperitoneally into a normal monkey. No symptoms developed and death occurred $5^1/_2$ weeks later. The autopsy failed to show any cause of death.

Another monkey was injected intracerebrally with infective mouse brain and became suddenly semi-comatose on the fifth day after injection. Mosquitoes (*Aëdes aegypti*) were given an infective feeding at this time. On the following day the monkey was in a dying condition and was sacrificed. The liver was extensively degenerated and occasionally fine acidophilic granules were noted in the nuclei, but there was no necrosis. A normal monkey was bitten by 26 of these mosquitoes, 17 days after their infective feeding but the monkey remained well and susceptible.

Serial Inoculations in Monkeys.—Serial subinoculations in four monkeys were carried out by injection of infective brain to brain, the first monkey of this series being injected with the brain of a monkey dying after intracerebral injection of mouse virus. The infective brain for these inoculations was taken from monkeys either at the moment of death or from animals which were sacrificed in a dying condition, that is, when coma had developed, accompanied by a subnormal temperature such as 94°F. In this series of animals, paralysis developed simultaneously on both sides. There was marked encephalitis but no necrosis of the liver. The first monkey injected with infective monkey brain showed the typical gross changes of yellow fever, namely, well-marked jaundice, extensive black hemorrhage into the stomach and a very yellow liver. It appeared as though the virulence of this mouse strain might be quickly restored to its normal condition by passage through monkeys, but on the contrary, the subsequent animals in this series did not show changes either in the gross or microscopically that were in any way comparable to the ordinary reaction of monkeys to yellow fever.

After passage through 4 monkeys this virus was injected into the brain of 6 mice and all died within 7 to 8 days.

Cross Immunity Tests.—In the course of these experiments five monkeys were inoculated into the brain with infective mouse brain and all died or were sacrificed in a dying condition, usually in 6 or 7 days after injection. This finding enables one to carry out cross immunity tests in two directions. Normal monkeys were injected intraperitoneally with the mouse strain and subsequently tested for immunity by the injection of infective

monkey blood. Conversely, monkeys immune to the typical virus of yellow fever were inoculated intracerebrally with mouse virus.

Intraperitoneal Injections.—Three normal monkeys were injected intraperitoneally with infective mouse brains and all remained apparently well. Two of these monkeys received 2 subsequent injections of mouse brain virus. These three monkeys were then injected with the typical monkey virus and two developed a sharp febrile reaction for one day only, after an incubation period of 3 and of 4 days, respectively. All three remained in excellent condition without any apparent signs of illness whereas 3 controls died typically of yellow fever.

Intracerebral Injections.—Four monkeys which had been immunized to the typical virus of yellow fever were injected intracerebrally with a suspension of infective mouse brain. Two of these animals remained perfectly well whereas their controls died, the sections of the brain showing a definite encephalitis. One of the immunes died on the 7th day after injection as a result of a bacterial abscess which developed at the site of inoculation in the brain. The control of this animal died at the beginning of the fourth day after injection. Histologically, degenerative changes were seen in the liver; the brain showed diffuse encephalitis complicated by a suppurative meningitis. It would appear that this animal was infected by the mouse virus and that death was hastened by an intercurrent meningitis. The fourth of these immune monkeys died 6 days after injection of a typical virus encephalitis. The liver was normal. It is noteworthy that this animal had received no immunizing injection for a period of one year and 8 and $^{2}/_{3}$ months.

For comparison with these observations, one monkey immune to typical yellow fever was injected into the brain with the standard virus using infective monkey blood. This animal remained well.

Pathology. Brain.—Histologically, the brains of the monkeys inoculated intracerebrally with the mouse strain of virus usually showed, in varying degree, definite evidence indicating a virus encephalitis. Occasionally nuclear inclusions were seen in the necrotic ganglion cells which were similar but not altogether identical with the inclusions found in infective mouse brains. These examinations were very kindly made by Dr. E. W. Goodpasture.

Liver.—Of five normal monkeys injected into the brain with mouse virus, none showed lesions of the liver comparable to the changes which occur in man or in monkeys dying of typical yellow fever. In one of these monkeys, a moderate amount of necrosis of the liver was found, and in another, the liver was normal. In the remaining three monkeys the liver showed moderate degenerative changes consistent with the earlier changes seen in yellow fever but by no means diagnostic and quite unlike the extensive necrosis seen in monkeys dying in the usual manner.

Of the four monkeys inoculated serially into the brain with infective monkey brain, the first showed marked fatty infiltration of the liver without any degeneration; in the next two, the liver was normal and in the fourth, there was degeneration without necrosis.

Of all the monkeys which received intracerebral injections, there was only one in which hemorrhage occurred into the stomach.

Two control monkeys were injected into the brain with the typical virus of yellow fever using 0.05 cc. of infective monkey blood. In both of these animals, definite necrosis occurred in the liver but no necrosis nor encephalitis was seen in sections of the brain.

Question of Immunization in Man.—Considerable evidence has gradually accumulated which indicates that a mild or even inapparent infection with yellow fever affords much more protection than any injection of killed virus. It is noteworthy that rhesus monkeys can be injected with impunity with the living mouse virus. The monkey appears to be even more susceptible than man to yellow fever and sooner or later the question may come up concerning the suitability of the mouse virus for protective inoculation in man under special circumstances. Such a suggestion would require careful control in two directions, namely, whether the injection of mouse virus in man would be free from danger and whether it would afford adequate protection. Unfortunately the accidental infections of laboratory workers with the mouse strain have not been fully recorded. The symptoms appear to be very mild. However, it is not established that the apparently mild infections without jaundice or albuminuria are free from danger.

Under any normal conditions in the endemic zones of yellow fever it is clear that reliance should be placed on measures for mosquito control. In recent years serious outbreaks of yellow fever have occurred under conditions which did not permit prompt control by anti-mosquito measures. Under special circumstances, it is conceivable that the inoculation of a strain of very low virulence might be attended with less hazard than is involved in the risk of spontaneous infection with a very potent strain of yellow fever.

Summary.—A strain of virus from a yellow fever monkey maintained by repeated passage from brain to brain of mice till it no longer infected monkeys (*M. rhesus*) by ordinary routes of injection was compared with the original strain of yellow fever maintained in monkeys and mosquitoes. Monkeys inoculated into the brain with a suspension of infective mouse brain developed regularly a typical fatal encephalitis but without characteristic changes in the liver. Blood of such animals proved non-infectious for monkeys, in two experiments, one test being made by direct inoculation and the other by means of mosquitoes.

Direct inoculation from brain to brain through a series of 4 monkeys failed to restore the virulence of the strain; obviously this experiment by

no means exhausts the possibilities of restoring the mouse strain of virus to its original condition in which it produces a fatal infection in monkeys after intraperitoneal injection with extensive necrosis of the liver. Such a restoration of virulence would be important both for theoretical and practical reasons.

Two monkeys injected intracerebrally with infective blood of monkeys died characteristically of yellow fever with necrosis of the liver and without the development of encephalitis.

Tests for cross immunity were carried out by injecting normal monkeys intraperitoneally with mouse virus and subsequently testing them with the typical virus of yellow fever. Also monkeys immunized to typical yellow fever were injected intracerebrally with infective mouse brains. Cross protection was very well marked though it was not entirely complete. The intraperitoneal injection of infective mouse brains proved to be a very convenient method for immunizing monkeys against a typical potent strain of yellow fever.

The results of these cross-immunity tests are entirely consistent with the interpretation that the virus in mice is yellow fever and there is no indication that it is contaminated by any secondary virus. However, the amount of data available at present is not overwhelming and there is no urgent need for drawing any altogether final conclusion. In the meantime, a more detailed investigation of these immunological findings is in progress.

* This work was supported by generous grants from the DeLamar Mobile Research Fund.

[1] Stokes, A., Bauer, J. H., and Hudson, N. P., *Amer. Jour. Trop. Med.*, **8**, 103 (1928).

[2] Theiler, M., *Ann. Trop. Med. and Parasit.*, **24**, 249 (1930). *Ibid.*, **25**, 69 (1931)

THE HEAT OF COMBUSTION OF METHYL ALCOHOL[1,2]

BY FREDERICK D. ROSSINI

NATIONAL BUREAU OF STANDARDS, WASHINGTON, D. C.

Read before the Academy, April 28, 1931

Within the past few years the heat of combustion of methyl alcohol has been the subject of much discussion on the part of those interested in the reaction involving the synthesis of methyl alcohol from carbon monoxide and hydrogen. Because of its industrial importance, the equilibrium conditions for this reaction have been studied by many investigators, whose data have been more or less concordant.

Several years ago Kelley[3] calculated the entropy of methyl alcohol from his calorimetric measurements of the heat capacity, and combining this with the heat of formation, calculated the free energy of formation of

methyl alcohol. However, when this value was combined[4] with the free energy of formation of carbon monoxide to give the free energy change for the reaction

$$CO + 2H_2 = CH_3OH, \qquad (1)$$

the resulting values for the equilibrium constant differed by a factor of about 10 from the experimentally measured ones. An inspection of the accuracy of the other factors in the calculation showed that this large discrepancy could be accounted for by a negative error of some $1^1/_2$ per cent in the value selected for the heat of combustion of methyl alcohol.

The reported values for the heat of combustion of methyl alcohol (liquid), at a constant pressure of 1 atmosphere, range from 170 to 173 kilocalories per mole—the former value by Favre and Silbermann[5] in 1852, and the latter calculated from the data of Thomsen[6] obtained in 1880. The usually selected "best" value has been that of Richards and Davis[7] who, in 1920, reported the value 170.8 kilocalories per mole.

In view of the discordant nature of the existing data, it seemed desirable to redetermine the heat of combustion of methyl alcohol.

The same calorimetric apparatus that was used in this laboratory to determine the heats of combustion[8] of hydrogen, methane and carbon monoxide was employed in the present work, and the same procedure followed.

In the present investigation methyl alcohol vapor was burned at constant pressure in a reaction vessel in the calorimeter. This was accomplished by saturating, at room temperature, a stream of air (free from water and carbon dioxide) with methyl alcohol vapor, and leading this gaseous mixture into the burner tube from which it emerged into an atmosphere of oxygen. The mixture was ignited by means of a spark, and the flame burned quietly at the burner tip. Most of the water formed was condensed to liquid in the reaction vessel. All of the carbon dioxide, and some water vapor, were carried out of the reaction vessel by the excess gas, which consisted of oxygen and nitrogen. On leaving the calorimeter the issuing gas passed first through an absorber containing "dehydrite" ($Mg(ClO_4)_2.3H_2O$) and phosphorus pentoxide, which absorbed the water, then through a second absorber containing "ascarite" (a sodium hydroxide-asbestos mixture) and phosphorus pentoxide, which absorbed the carbon dioxide, and finally through a guard tube. The amount of reaction was determined from the mass of carbon dioxide absorbed. The increase in weight of the absorber was corrected to vacuum to give the true mass of the carbon dioxide of which 44.000 g. was taken as equivalent to 1 mole of methyl alcohol.

The thermal effect produced in the calorimeter by the energy of reaction was duplicated as nearly as possible in experiments with electrical energy.

In this manner, the heat evolved in the combustion of methyl alcohol to form a measured mass of carbon dioxide was determined by measuring the quantity of electrical energy which was required to produce the same amount of heat.

The calorimetric data of the combustion experiments were corrected for (1) the energy introduced by the sparking operation, (2) the energy required to bring the entering gases to the average temperature of the calorimeter, and (3) the energy involved in the vaporization of that mass of water not condensed to liquid. In all the experiments, the temperature rise of about 3 degrees occurred in about 30 minutes. The relation between the calorimeter temperature and the time was similar to that in the experiments on hydrogen and oxygen.[8]

The methyl alcohol used in the present work was purified, from the best synthetic material available, by distillation in a 30 plate bubbling cap column.[9] Only the middle portion of the distillate was used. As reported in another investigation[10] on the thermal properties of methyl alcohol, this sample had a density ($d_4^{20} = 0.79133$) practically identical with the "best" value given by the International Critical Tables.[11] A series of six combustion analyses to determine the ratio of carbon to hydrogen in this alcohol was made by passing the methyl alcohol vapor through copper oxide at 600–800°C., and absorbing the water and carbon dioxide formed in "dehydrite" and "ascarite," respectively.

These experiments gave for the ratio, $\frac{2(\text{moles } CO_2)}{(\text{moles } H_2O)}$, the value 1.0000 ± 0.0003.

Tests showed that the process of saturating the carrying air stream with methyl alcohol caused no oxidation of the alcohol. Examination of the products of the reaction as carried out in the calorimeter showed the

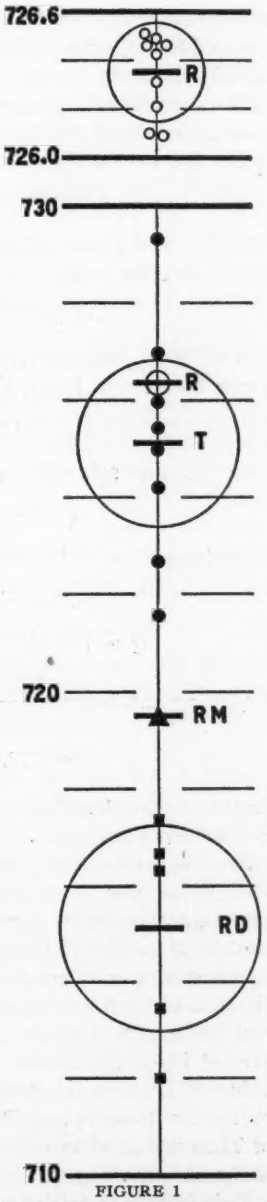


FIGURE 1
Plot of the data on the heat of combustion of methyl alcohol.

presence of only a negligible amount of formaldehyde (1 mole per 800,000 moles CH_3OH) and of nitrogen oxides (1 mole per 40,000 to 60,000 moles CH_3OH). Tests by two different methods, which were sensitive to 0.003 and 0.001%, respectively, showed no carbon monoxide in the gaseous products of combustion.

For the heat evolved in the reaction

$$CH_3OH_{(g)} + 3/2O_{2(g)} = CO_{2(g)} + 2H_2O_{(l)} \quad (2)$$

at 25°C., and a constant pressure of 1 atmosphere, the data of nine experiments give the value

$$Q = 763.77 \pm 0.20 \text{ int. kilojoules per mole.} \quad (3)$$

The $CH_3OH_{(g)}$ had a partial pressure about equal to its saturation pressure. From the data of Fiock, Ginnings, and Holton[10] one finds for

$$CH_3OH_{(l)} = CH_3OH_{(g)} \quad (4)$$

at 25°C., and saturation pressure, that

$$Q = -37.43 \text{ int. kilojoules per mole.} \quad (5)$$

Combining these data gives for the reaction

$$CH_3OH_{(l)} + 3/2\ O_{2(g)} = CO_{2(g)} + 2H_2O_{(l)} \quad (6)$$

$$Q = 726.34 \pm 0.20 \text{ int. kilojoules per mole} \quad (7)$$

Or, using the factor $\frac{1.0004}{4.185}$,

$$Q = 173.63 \pm 0.05 \text{ kg.-cal.}_{15} \text{ per mole.} \quad (8)$$

In the above calculation the heat of mixing air and methyl alcohol vapor is considered negligible.

For comparison, the recomputed data of Thomsen,[6] and those of Richards and Davis[7] and Roth and Müller,[12] are assembled together with the present data in figure 1, where the ordinate scale gives the heat of combustion of methyl alcohol (liquid), at 25°C. and a constant pressure of 1 atmosphere, in international kilojoules per mole. The points designate the data of the following investigators: ■, Richards and Davis; ▲, Roth and Müller; ●, Thomsen; ○, Rossini. The individual values from the present investigation are plotted in the upper part of the figure. The value obtained in the present work is 1.5 per cent higher than that reported by Richards and Davis,[7] while the value computed from the data of Thomsen[6] is in agreement with the present result within the assigned limits of error.

This work was carried out under the direction of E. W. Washburn to

whom the author is greatly indebted for his deep interest and valued suggestions. Grateful acknowledgment is made to J. H. Bruun for purifying the methyl alcohol, and to the Gas Chemistry Section of this Bureau for developing the flame technic and making the carbon monoxide tests.

[1] Publication approved by the Director of the Bureau of Standards, U. S. Department of Commerce.

[2] A complete account of this work will appear in the *Bur. Stds. J. Res.*

[3] Kelley, *J. Am. Chem. Soc.*, **51**, 180 (1929).

[4] Kelley, *Ind. Eng. Chem.*, **21**, 353 (1929).

[5] Favre and Silbermann, *Ann. chim. phys.*, **34**, 357 (1852).

[6] Thomsen, "Thermochemische Untersuchungen," **4**, p. 157, Barth, Leipzig (1886).

[7] Richards and Davis, *J. Am. Chem. Soc.*, **42**, 1599 (1920).

[8] Rossini, *Proc. Nat. Acad. Sci.*, **16**, 694 (1930); *Bur. Stds. J. Res.*, **6**, 1, 36 (1931).

[9] Bruun, *Ind. Eng. Chem., Anal. Ed.*, **1**, 212 (1929).

[10] Fiock, Ginnings and Holton, *Bur. Stds. J. Res.*, **6**, 881 (1931).

[11] International Critical Tables, **3**, p. 27, McGraw-Hill Book Co., Inc., New York, 1928.

[12] Roth and Müller, Landolt-Börnstein-Roth-Scheel Tabellen, p. 868, Springer Berlin, 1927.

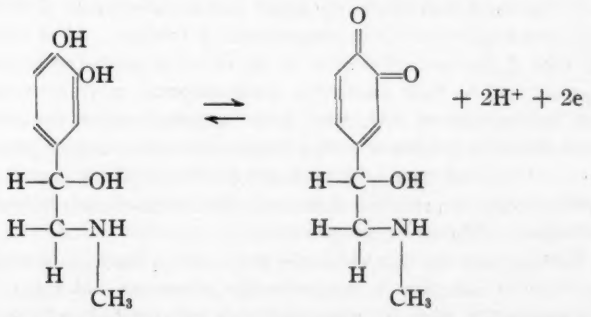
A POTENTIOMETRIC STUDY OF EPINEPHRINE

BY ERIC G. BALL* AND W. MANSFIELD CLARK

DEPARTMENT OF PHYSIOLOGICAL CHEMISTRY, THE JOHNS HOPKINS UNIVERSITY, SCHOOL OF MEDICINE

Communicated May 2, 1931

Epinephrine is a derivative of catechol. As such it should yield, as a first stage of oxidation, the corresponding orthoquinone



By analogy such a system should be reversible and should establish at a noble metal electrode a potential indicative of the equilibrium state.

However, the ordinary methods are inapplicable because the oxidant decomposes with great rapidity.

With this system and also with catechol we had previously employed the procedure published by Fieser.[1] This consists in measuring the potentials of the system at known intervals of time after establishing the mixture and extrapolating to zero time or in other ways resolving the potential of the true system from the drifts due to removal of components and to accommodations of individual electrodes.

Since the results were not entirely satisfactory in the case of catechol and were certainly very inaccurate in the case of epinephrine we then employed a method similar to that introduced by Hartridge and Roughton[2] for measuring the rate of oxygen uptake by erythrocytes. The details of the apparatus will be described in a more complete account of the experiments.

The method is as follows. A buffered solution of the reductant and a buffered solution of an oxidizing agent are driven into a mixing chamber. The mixture flowing from this chamber over a series of electrodes bathes each electrode with a solution of presumably fixed composition. The proportion of oxidant and reductant can be calculated from the concentrations and the relative rates of flow of the two solutions. Tests of the completeness of mixture can be judged from uniformity in potentials observed in cases where the two solutions contain respectively a preformed stable oxidant and a preformed stable reductant, e.g., quinone and hydroquinone. A complete resolution of effects due to slow interaction of reductant and oxidizing agent (e.g., ceric salt) and those due to decomposition of oxidant has not been attained. However, the known volume spacing of electrodes and the known rates of flow enable one to plot a curve of potential against time. Such a curve shows with catechol in acetate buffer a very rapid rise of potential presumably due to the interaction of reductant and oxidizing agent and a subsequent slower fall of potential presumably due to decomposition of oxidant. It is tentatively assumed that if this second section of the curve is projected to zero time the potential will be that which the undecomposed mixture would have were the interaction of reductant and oxidizing agent instantaneous. Since each electrode is bathed with a solution of a composition presumably constant at this point, and is so bathed for a considerable period of time, no allowance need be made for slow accommodation of the electrode itself as is necessary in the ordinary procedure.

After flowing past the last electrode the solution forms a flowing liquid junction with a solution saturated with potassium chloride. This is in direct connection with a "saturated KCl calomel half-cell."

In these preliminary measurements no accurate temperature control was used but the apparatus was placed in a room of fairly constant temperature.

Table 1 gives the results with the catechol system. E_0' is the calculated potential of an equimolecular mixture of the oxidant and reductant at the pH value indicated.

TABLE 1

CATECHOL (0.004 *N*) OXIDIZED WITH CERIC SULFATE (0.002 *N*) IN ACETATE BUFFER pH 4.40. TEMP. 23.0°C.

REDUCTION PER CENT	$0.03006 \times \log \frac{[S_r]}{[S_0]}$	E_h OBSERVED	E_0'
15.3	−0.0223	0.5625	0.5402
22.6	−0.0161	0.5548	0.5387
36.6	−0.0072	0.5468	0.5396
47.8	−0.0011	0.5413	0.5402
57.5	+0.0039	0.5351	0.5390
68.7	0.0103	0.5292	0.5395
81.7	0.0195	0.5200	0.5395
91.0	0.0302	0.5087	0.5389
			Av. 0.5395

If we assume that the relation between pH and E_0' in the acid region is defined by $\frac{\Delta E_0'}{\Delta \text{pH}} = -0.0587$, the normal potential (at pH = 0) would be 0.798 at 23°C. Fieser[3] reports 0.794 at 25°C.

Table 2 contains the results of an experiment with epinephrine. The sample used in this instance was the bitartrate from the H. A. Metz Laboratories.

TABLE 2

EPINEPHRINE BITARTRATE (0.004 *N*) OXIDIZED WITH CERIC SULFATE IN ACETATE BUFFER pH 4.40. TEMP. 24.0°C.

REDUCTION PER CENT	$0.03006 \times \log \frac{[S_r]}{[S_0]}$	E_h OBSERVED	E_0'
36.09	−0.008	0.538	0.530
57.94	+0.004	0.526	0.530
67.80	0.010	0.519	0.529
77.14	0.016	0.512	0.528
			Av. 0.529

A direct comparison between tables 1 and 2 is possible and reveals the epinephrine system as having a characteristic potential about ten millivolts negative to that of the closely related catechol system.

The rate of decomposition of the oxidant of epinephrine is about one hundred-fold that of orthoquinone at pH 4.4. About 98% of the oxidized epinephrine disappears in 15 seconds. This rate increases rapidly with increase of pH and decreases with decrease of pH. At pH 0.03 only a slight drift of potential was observed. At pH 7.45 the potential after the first half-second drifted 60–100 millivolts per second so that at the end of two seconds approximately 100% of the oxidant had disappeared.

In table 3 are summarized a series of preliminary experiments with different preparations of epinephrine at different pH values. Further work will be necessary to define clearly the relation of E_0' to pH and particularly to establish beyond question the characteristics of the system in the biologically important region of pH.

TABLE 3

E_0' VALUES OF SEVERAL PREPARATIONS OF EPINEPHRINE. TEMP. 24.0°C. ± 1.0°

PREPARATION	OXIDIZING AGENT	pH	E_0'	CALCULATED NORMAL POTENTIAL ASSUMING A 0.0589 SLOPE (24.0°C.)
Student I	Ceric sulfate	4.40	+0.525	0.784
Student II	Ceric sulfate	4.40	0.519	0.778
Dr. Abel*	Ceric sulfate	4.40	0.523	0.782
Metz	Ceric sulfate	4.40	0.529	0.788
Metz	Ceric sulfate	0.03	0.792	0.794
Metz	Ferricyanide	6.66	0.395	0.787
Metz	Ferricyanide	7.45	0.343	0.782

* A sample kindly supplied by Dr. Abel. Prepared from bufo agua.

Nevertheless three rather important conclusions may be drawn.

1. The characteristic potentials of the epinephrine system are close to those of the related catechol system. This is in agreement with the usual rule that closely related systems have approximately the same potentials.

2. The inherent instability of orthoquinone is enhanced by the presence of a side chain with ionizable groups. This is in agreement with preliminary results we have obtained with quinones containing side chains with ionizable groups, e.g., the oxidant of homogentisic acid.

3. Cannan, Cohen and Clark[4] reported the experimental fact that biological material, in general, establishes an electrode potential and acts upon defined oxidation-reduction indicators in a manner suggesting the maintenance by cells of a reduction intensity distinctly negative to the region which characterizes the indophenols.

This conclusion has been confirmed by numerous investigators. But to establish its validity as well as its significance the fact must be shown to be consistent with the found states of naturally occurring oxidation-reduction systems. Conant,[5] and Conant and Fieser[6] have determined the potential of the hemoglobin–methemoglobin system. This and the chemically undefined "reduction potential" of the tissues are such that one can say the reducing tendency of the tissues tends to keep the hemoglobin from oxidizing to methemoglobin and ready to perform its function as a carrier of chemically inert oxygen. The reduction potential of the epinephrine system and the chemically undefined "reduction potentials" of the tissues are such that the cells tend to protect this hormone from an oxidation which would result in an extremely rapid destruction.

* NATIONAL RESEARCH COUNCIL FELLOW in Medicine, 1929–1930.

[1] Fieser, Louis F., *J. Am. Chem. Soc.*, **52**, 4915, 1930.

[2] Hartridge, H., and Roughton, F. J. W., *Proc. Roy. Soc., Ser. A.*, **104**, 376, 1923.

[3] Fieser, Louis F., *J. Am. Chem. Soc.*, **53**, 793, 1931.

[4] Cannan, R. K., Cohen, B., and Clark, W. M., *Supplement No. 55, Public Health Reports*, 1926.

[5] Conant, J. B., *J. Biol. Chem.*, **57**, 401, 1923.

[6] Conant, J. B., and Fieser, L. F., *Ibid.*, **62**, 595, 1925.

THE DISTRIBUTION OF STRESSES IN WELDED AND RIVETED CONNECTIONS

BY WILLIAM HOVGAARD

DEPARTMENT OF NAVAL ARCHITECTURE, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Read before the Academy, April 27, 1931

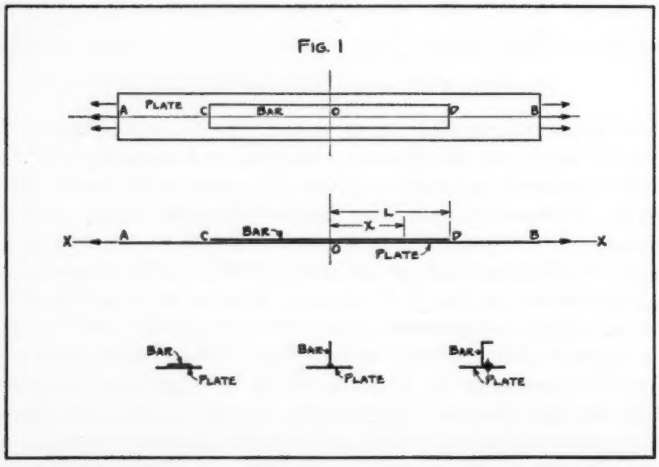
A method for determining the stresses in a line of weld or rivets subject to shearing in its own direction was described in a previous paper.[1] We shall here explain its application to various important problems occurring in Naval Architecture and Civil Engineering, and briefly summarize the results of tests recently made at the Massachusetts Institute of Technology. The fundamentals of the method will be first recapitulated.

A girder of limited length is attached to a major structure which is subject to tension or compression in a direction parallel with the girder. It is required to determine the stresses in the connection, whether welded or riveted, as well as in the adjacent part of the structure. We consider in particular the elementary case of a bar or narrow plate attached to a plate of greater length and width, and refer to the minor structure as the "bar" and to the major structure as the "plate." Figure 1 shows such a plate and various forms of "bar."

When the plate is subject to elongation due to a simple uniform pull at the ends, the bar will be forced to follow due to shearing in the weld or rivets which connect it to the plate; but it is clear that the bar will offer resistance and that there will be a certain elastic creeping of it relative to the plate. Thus every section of the bar will suffer a displacement relative to the corresponding section of the plate, but by symmetry this displacement must be zero at the middle and it is obvious that it will be a maximum at the ends of the bar. Now the assumption is made that the shearing stress in the weld or rivets is proportional to the displacement, or we may say that the displacement at any transverse section is equal to the shearing stress at that section multiplied by a constant, which we call the "displacement coefficient" and denote by μ. The value of μ

is not the same in different structures, and can only be determined by experiments.

The theoretical solution of the problem furnishes a very simple expression for the shearing stresses in the connection, as well as for the average tensile stresses in the bar and in the plate. Across any transverse section the total pull on the bar must be equal to the total shearing force exerted by the connection outside that section while the pull in the plate is reduced by the same amount. These conditions furnish two equations, while a third is obtained from the fact that the creeping or displacement of the bar relative to the plate at any section must be equal to the difference in elongation of the plate and the bar reckoned from the middle up to that section. The form of the function expressing the shearing stress is

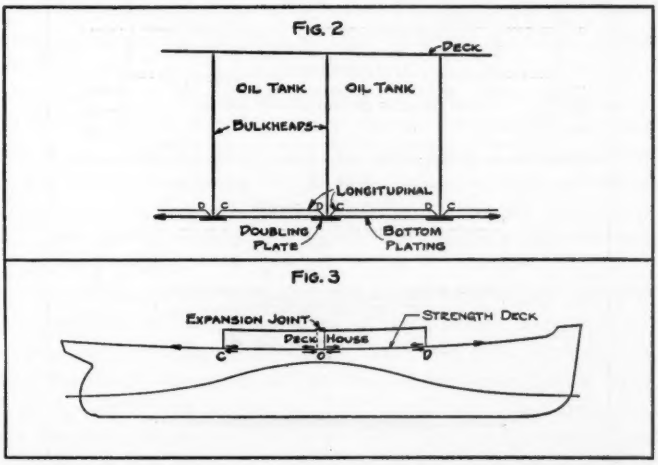


determined by the method of variation, as explained in the paper referred to above, so as to make the elastic work stored up in the entire structure a minimum. In forming the integral which expresses the elastic work it is assumed that the stresses across the bar as well as those across the plate are uniform and equal to the mean value of the actual stresses in each member at any given cross-section. Hence the displacements of the bar relative to the plate calculated on this assumption are also average displacements. Actually, of course, the stresses vary across a section.

So far only a small number of tests have been made, but they seem to confirm the constancy of μ and show conclusively that there is a marked concentration of shearing stresses at the ends of the bar as predicted by the theory. The intensity of these stresses is believed to be much greater than ordinarily supposed.

We shall now describe some of the applications of the theory.

1. *Longitudinal Girders in Ships.*—Ordinarily longitudinals fitted in the bottom of ships are continuous and extend over the greater part of the length of the ships. Since the longitudinal bending moments acting on the ship are small or moderate at the ends, the bending stresses in the bottom plating are small in these parts and the shearing stresses in the rivets which connect the longitudinals to the plating are small. The longitudinals form in this case an integral part of the ship's structure and practically share the strains and stresses with the bottom plating throughout their length. In certain oil tankers, however, short intercostal longitudinal girders are fitted to the bottom plating between the transverse bulkheads, and no connection between these longitudinals exists through



the bulkheads. See figure 2. It may be expected, according to the theory, that when such a ship is exposed to great bending moments, as when she is forced up against a heavy sea, excessive shearing stresses are liable to occur at the ends of the longitudinals, C and D, especially at amidships, causing leaky rivets Serious defects of this nature have actually occurred in several cases, showing the undesirability of attaching such short detached girders to a major strained structure. It was in fact an analysis of the leakages that occurred in certain oil tankers of this construction which led the author of this paper to develop the theory here presented.

2. *Deckhouses.*—When deckhouses are fitted on the strength deck of a ship near the half-length where great bending stresses are apt to occur, they will be in a position similar to that of a bar connected to a strained plate. See figure 3. So long as the deckhouse is left intact, no shearing

stresses will exist at the middle, O, but excessive shearing stresses are liable to be found at the lower corners C and D, and may cause fracture of the adjoining deck plating as well as at the corners of doors and windows in the sides of the deckhouse. So serious and common has been this defect that in many ships deckhouses have been cut in two or more parts, and

FIG. 4

HATCH OPENINGS IN A STRENGTH DECK

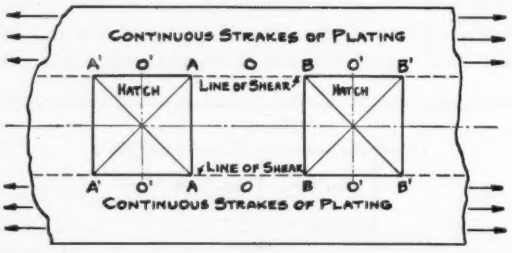
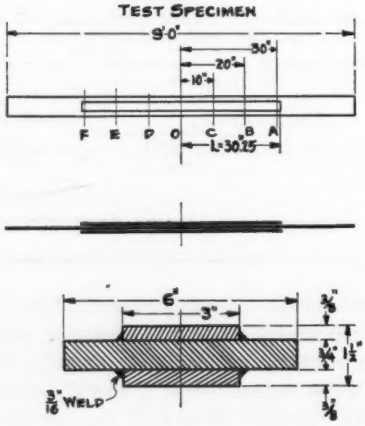


FIG. 5

TEST SPECIMEN



fitted with so-called "expansion joints" in the hope that the stresses should be thereby relieved. Such joints did not, however, relieve the stresses at the corners of the deckhouses and, in fact, in many cases aggravated the difficulties by multiplying the points of discontinuity. Attention was fixed on the stresses due to bending in the top of the deckhouse, while the stresses at the lower corners remained unexplained. It is not as-

serted, on the other hand, that expansion joints are otherwise useless, since they may prevent the direct stresses from being excessive in the top of the deckhouse when this structure is of great height.

3. *Large Hatch Openings.*—In most ships large hatch openings are found in the strength deck, as shown in figure 4, destroying the continuity of the central strakes of plating while leaving the outer strakes, the so-called deck stringers, intact. In many ships fracture of the deck plating has occurred at the corners of such hatches, in some cases, as recently in the *SS Majestic* and *SS Leviathan* with very serious consequences. It can be seen that fundamentally these defects, popularly explained as being due to discontinuities, are due to the same cause as the breakdown at the corners of deckhouses. Suppose that the stringer plates outside the hatches on both sides elongate due to bending of the ship, then the plating between the hatches, which would otherwise be practically inert, will be forced to follow, but this it can do only through the medium of shearing stresses along the line AB. At the points O, midway between the hatch corners A and B, there will be no shearing stress by symmetry, but as shown by the theory there will be a marked concentration of shearing stresses at the hatch corners A and B. This effect is augmented by the shearing stresses which occur for the same reason at these corners due to the presence of the deep longitudinal coaming plates on the sides of the hatches, AA' and BB', which are likewise, each of them, required to elongate with the deck plating.

4. *Cover Plates on the Flanges of Bridge Girders.*—Such plates are in the same position as the typical bar described above, being riveted to a major strained structure. Here again there will be a concentration of shearing stresses in the rivets at the ends of the cover plates, while the rivets at the middle will be inactive. It is to be noted that these shearing stresses are not a function of the shear forces which act across the girder; in fact, they would exist even if there were no shear forces and only a pure bending couple was acting.

APPENDIX

Experiments Carried Out at the Massachusetts Institute of Technology

The principal object of these experiments was to verify the assumed constancy of the displacement coefficient and to determine its value in various cases.

An experimental specimen is shown diagrammatically in figure 5, on which figured dimensions are given. It consists of a plate strip on each side of which is fitted a thinner strip, called the "bar," welded to the plate by continuous fillet welds.

The displacement of points on the bar and on the plate close to the

weld was measured directly at both ends, A and F, and at two intermediate points C, B, and DE on each side of the middle, O. Displacements were also obtained less directly from measurements of strains along the center line of the bars and along other lines on the bars and the plate parallel with the center line. The strains were measured partly with 20-inch and partly with 2-inch Berry gages, the latter being applied at five stations on which seven measurements were made on each side of the specimen.

The specimen was placed in a 400,000-lb. Emery testing machine and subjected to tensions increasing from a small initial load up to about 200,000 lbs., corresponding to average stresses in the plate outside the bar of about 45,000 lbs. per sq. in., while the total section of plate and bar was stressed to an average of about 30,000 lbs. per sq. in. The machine was just able to hold this load, but when going somewhat higher, the elongation increased rapidly and at 250,000 lbs. the plate fractured at the lower end through the holes drilled for the measuring devices. It appeared that at a load of 105,000 lbs. the entire specimen was still within the elastic limit, and the analysis was carried out for this load.

We give here the definition of the symbols and state the formulas that were used in the analysis. The derivation of these formulas is given in the PROCEEDINGS referred to above.

$2L$ = total length of bar
A = sectional area of plate
a = sectional area of bar
α = sectional area of weld per unit length of bar
q_x = shearing stress in weld at the point x
$\underset{\sim}{p}_x$ = average tensile stress in plate across a section at x
$\bar{p}_x$ = average tensile stress in bar across a section at x
μq_x = average linear displacement of a section of the bar relative to the corresponding section of the plate at x.

The shearing stress in the weld at the end of the bar is given by:

$$q_L = \frac{p \tanh mL}{m\mu E} \tag{1}$$

where

$$m = \sqrt{\frac{\alpha(A + a)}{aA\mu E}}. \tag{2}$$

The shearing stress at any point:

$$q_x = \frac{q_L \sinh mx}{\sinh mL} = \frac{p}{m\mu E} \cdot \frac{\sinh mx}{\cosh mL}. \tag{3}$$

The tensile stress in the bar at any section:

$$\bar{p}_x = \frac{pA}{A + a} \frac{\cosh mL - \cosh mx}{\cosh mL}. \quad (4)$$

The stress in the plate:

$$p_x = p - \frac{a}{A} \bar{p}_x. \quad (5)$$

In the present case:

$A = 4.50$ sq. in. $\quad L = 30.25$ in.

$a = 2.25$ sq. in. $\quad$ Total shearing area in weld on half length = 16.0 sq. in.

$\alpha = 0.530$ sq. in. $\quad E = 30 \times 10^6$ lbs. per sq. in.

The observed displacements were averaged for three sections, placed, respectively, 10 in. and 20 in. from the middle and at the ends, each average being a mean of eight readings.

For a load of $P = 105{,}000$ lbs. the stress in the plate outside the bar was:

$$p = \frac{105{,}000}{4.5} = 23{,}300 \text{ lbs. per sq. in.}$$

The displacement at the end of the bar was 12.5×10^{-4} in. and this should be equal to μq_L. Hence from (1):

$$12.5 \times 10^{-4} = \mu \frac{p \tanh mL}{m\mu E}.$$

As will be seen, $\tanh mL = 1$ in this case, so that:

$$m = \frac{23{,}000}{12.5 \times 10^{-4} \times 30 \times 10^6} = 0.62$$

$$m^2 = 0.38, \quad mL = 0.62 \times 30.25 = 18.8.$$

From (2):

$$\mu = \frac{\alpha(A + a)}{aAEm^2} = \frac{0.53 \times 6.75}{10.13 \times 30 \times 10^6 \times 0.38} = 0.309 \times 10^{-7}.$$

We adopt the value $\mu = 0.31 \times 10^{-7}$ and find from (1):

$$q_L = \frac{23{,}300}{0.62 \times 0.31 \times 10^{-7} \times 30 \times 10^6} = 40{,}000 \text{ lbs. per sq. in.}$$

By an analysis of the data for $P = 80{,}000$ lbs. approximately the same value of μ was obtained, and, as stated in a previous paper,[2] certain experiments carried out at the University of Pittsburgh gave about the same result.

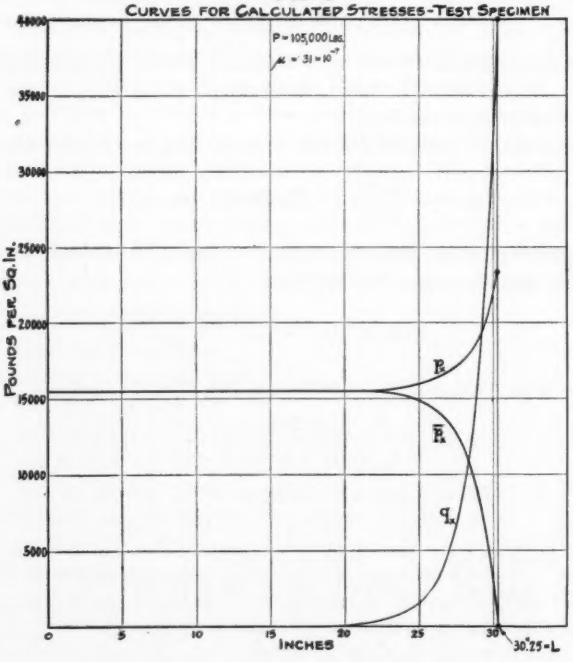
From formula (3) the values of the shearing stress q_x was determined for various values of x and plotted in a curve on figure 6. The stress is seen to be zero up to a distance of about 20 inches from the middle point of the bar, after which it rises very steeply, showing that the welds are

active only for a length of 10 inches at each end. It follows that over the middle portion of the bar, for a length of about 40 inches, the continuous weld could be replaced by short intermittent welds, just sufficient to prevent buckling under compressive load.

The tensile stresses in the bar and in the plate were calculated from formulas (4) and (5), respectively, and plotted in curves on figure 6. The stress in the bar is equal to the stress in the plate, 15,500 lbs. per sq. in., for 20 inches on each side of the middle, after which $\bar{p}_x$ falls off abruptly

FIG. 6

CURVES FOR CALCULATED STRESSES-TEST SPECIMEN



to zero at the end of the bar and p_x rises to 23,300 lbs. per sq. in. at that point.

It must be noted that since the displacements of the bars relative to the plate could not be measured along the central portion of the plate, which was covered by the bars, it is likely that the value of μ here obtained is somewhat too small. If that is the case, the value obtained for the maximum shearing stress is too high and the curve for q_x should be less steep.

The displacements observed directly at the ends of the bars and at the two intermediate stations at points located close to the weld gave the

best result, while the displacements obtained indirectly from strain measurements were unsatisfactory.

As stated above, the direct measurements made close to the welds showed a displacement at the end of the bar of 12.5×10^{-4} inches, but already at less than ten inches from the end it was zero, conforming in this respect closely with the curve for the shearing stress q_x.

[1] These PROCEEDINGS, November, pp. 667–673, 1930.

[2] *Ibid.*, Nov., p. 678, 1930.

THE RELATIVE EFFECTIVENESS OF SPECTRAL RADIATION FOR THE VISION OF THE SUN-FISH, LEPOMIS

BY HARRY GRUNDFEST*

LABORATORY OF BIOPHYSICS, COLUMBIA UNIVERSITY

Communicated April 24, 1931

1. This paper presents the results of an investigation on the visibility of spectral lights for the sun-fish, Lepomis. A new method has been developed by means of which measurements have been obtained on individual animals, both for dim and for bright lights.

Determinations of the relative efficiencies of very dim spectral lights for the eye of a fish were carried out for the purpose of testing the hypothesis that visual purple is the photosensitive substance mediating dim vision. This hypothesis rests chiefly on the fact that there is a close correlation between the relative amounts of spectral lights absorbed by human visual purple, their relative efficiencies in bleaching it, and their relative effectiveness in producing a liminal sensation of brightness in the human eye. This correlation was first pointed out by Koenig (1894),[8] later by Trendelenburg (1905)[17] and recently, with greater accuracy by Hecht and Williams (1922).[6] The Grotthus-Draper law of photochemistry which states that lights are effective only to the extent that they are absorbed makes such correlation a necessary condition if visual purple is to be the sensitive substance concerned in dim vision.

Through the work of Koettgen and Abelsdorff (1896)[10] we know that there are two types of visual purples as judged by their absorption spectra. One type is found in mammals, birds and amphibia and has a maximum absorption near the wave-length 500 mμ. The other type of visual purple is found in fish. Eleven species examined by Koettgen and Abelsdorff all have a visual purple whose maximum absorption is about 540 mμ. If it is true that visual purple is the sensitive substance for vision at dim illuminations, then the relative effectiveness of spectral lights in producing

a constant liminal effect in the eye of fish should be different from that for other vertebrates.

Measurements made in the course of this work with the visual purple of the sun-fish, Lepomis, have shown that its maximum absorption is also at 540 mμ. Therefore determinations of the visibility curve of Lepomis at very low illuminations make a test of the above hypothesis possible.

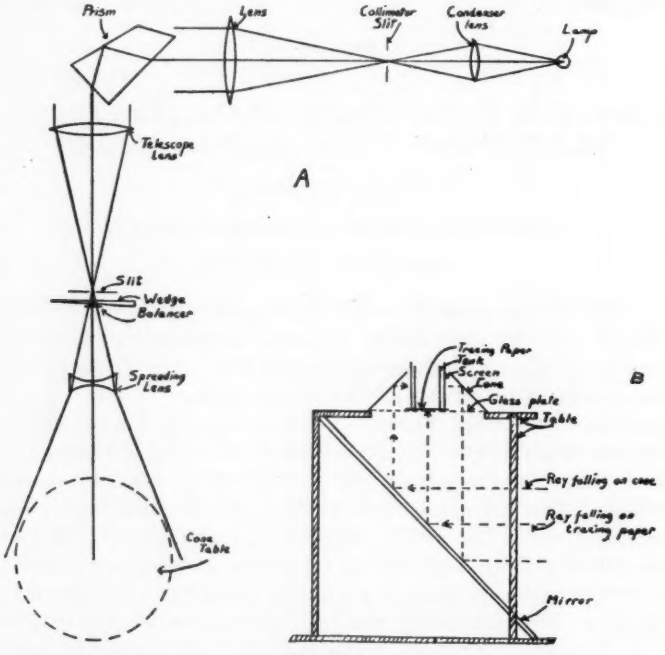


FIGURE 1

Arrangement of the apparatus. *A* is a top view showing the general disposition of the set-up. *B* is a vertical section of the box holding the mirror, screen, tank and animal.

2. The method which has been developed to measure the visibility function of fish takes advantage of the fact that a fish follows, with its body, a movement in its visual field. This reaction was first studied by Lyon (1904)[12] who showed that it was a photic and not, as previously supposed, a rheotropic response. The unfortunate name "rheotropic reaction" has remained, however, and is still used for this phenomenon.

The common sun-fish, Lepomis, is particularly accurate in its response to a moving visual stimulus and has therefore been employed in these

experiments. It is placed (Fig. 1) in a cylindrical glass jar on a glass-topped table. Around the outside of this tank is a cylindrical screen composed of equal and alternate vertical bars and spaces. A beam of light which is sent up through the glass table top is reflected through this screen and toward the animal in the tank by a hollow, truncate 45° cone made of plaster of Paris. The bottom of the glass tank is covered with translucent paper so that the portion of the light beam which falls on it produces an even and dimly illuminated background against which the animal can be observed.

When the illumination is sufficiently bright, rotation of the screen at the rate of about 1 cm./sec. causes the fish, which is generally quite motionless, to swim suddenly in the direction of the rotation. When the rotation of the screen is reversed, the fish correspondingly reverses its direction. The animal follows in this way continued reversals in rotation of the screen for minutes at a time.

On decreasing the illumination, a point is reached at which no response is produced. The minimum intensity at which the fish still responds varies with the size of the bars and spaces which make up the screen. Indeed, this method has been used by Hecht and Wolf (1929)[7] for measuring the relation between intensity and visual acuity in the honey bee and for Drosophila by Hecht and Wald (unpublished). The intensity, or more precisely, the energy content of the beam at which this reaction is just elicited also varies with the wave-length of the light. Thus, the efficiency of spectral lights for the eye of the sun-fish can be determined by measuring, for various spectral bands, the least amounts of energy which will just produce an oriented reaction of the animal in response to the movement of a given screen. In these experiments the bars of the screen are 2 mm. wide, so that the minimum illuminations necessary to obtain a rheotropic reaction are very low. The curves determined at these low intensities, which are of the order of 10^{-6} millilamberts, describe, therefore, the relative efficiency of spectral lights for the dim vision of the sun-fish.

Spectral lights are obtained from a Hilger constant deviation spectrometer. The relative energy content of each band used is determined with a Hilger thermopile and a Leeds and Northrup galvanometer. The amount of energy in the beam is controlled by means of a balanced neutral wedge whose transmission at each setting is accurately known. Experimental procedure consists in a determination of the wedge setting at which a response can just be elicited with each spectral band, from an animal that has been dark adapted for two hours.

3. Thirteen animals have been measured, many of them several times, on different occasions. It was found that the wave-length of maximum efficiency for the sun-fish lies between 535–545 mμ. This is in agreement

with the position of the maximum of the visual purple absorption curve (540 mμ), and indicates that the hypothesis relating visual purple with dim vision is correct. The shape of the visibility curve, however, does not even approximate the shape of the visual purple absorption curve. The former is much narrower and drops away from the maximum asymmetrically. The violet portion of the curve is variable among different individuals, whereas the other side is relatively constant. A typical result is shown in figure 2.

The clue to an explanation of this difference between the visibility and the absorption data is provided by one exceptional individual which had different visibility curves at different times. The first measurements for this fish coincided on the violet side with the similar portion of the visual purple absorption curve. The side toward the red, however, still agreed with the corresponding section obtained for all the other animals. In figure 3, the points represent the data. In the lower half of the figure, the full line is the reciprocal of the visual purple absorption curve. It is the theoretical visibility curve. The broken line is the red portion of the curve obtained from the other animals. Clearly, the data for this individual represent, on the violet side, a very close approach to the expected visibility curve, while on the red side the deviation found in other animals still persists. Three days after these measurements, this animal gave a visibility curve intermediate in shape between its first curve and one found with other individuals and in two weeks the visibility function had become quite similar to that usually found and is given in the upper half of figure 3. These data indicate that the shape of the visibility curve tends toward that demanded by the idea that visual purple is concerned in dim vision, but that other factors probably influence the amounts of spectral energies required by the sun-fish for liminal vision.

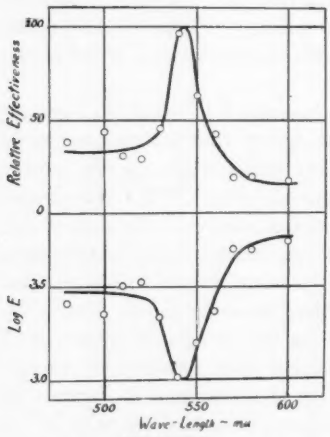


FIGURE 2

Visibility function at the lowest illuminations. The data are for animal 2NS and are entirely typical.

This discrepancy between visual data and visual purple absorption data may be accounted for on the assumption that there are present, besides visual purple, two other light-absorbing pigments which are not light sensitive and which act as filters. Such pigments are known to occur in the eyes of vertebrates, for example, the macular pigment in man,

and certain colored oil globules in birds. Among fish there have been found several eye pigments, of which carotin, guanin and melanin are most common (Cunningham and MacMunn, 1893;[2] J. Verne 1926[19]). Although it is generally stated that these pigments migrate behind the light sensitive elements in the dark-adapted retina, it is possible that they do not move back completely and so absorb some of the light which reaches the retina (cf. illustration in Wunder, 1925;[18] figure 16, p. 36). Carotin has absorption bands approximately where the filter affecting the violet portion of the visibility curve would absorb most. It is also significant that the concentration of carotin varies easily with the physiological condition of the animal (Palmer, 1922).[15] Such changes in the concentration of the pigment would account for the variability in the violet region. The presence of such pigments may also explain the results of Murr (1927),[14] on the cat, and of Dieter (1929)[3] on man.

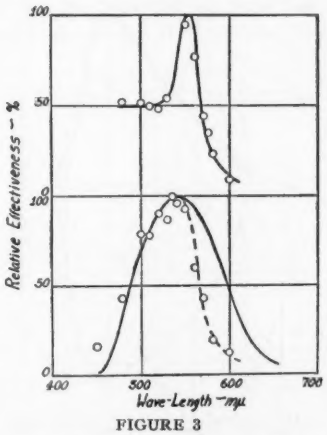


FIGURE 3

The visibility curve of animal 8 at two separate times. The data in the upper half of the figure were secured three weeks after those in the lower. In the lower half the continuous line is the absorption spectrum of fish visual purple according to Koettgen and Abelsdorff, while the broken line is the right half of the usual low intensity visibility curve.

4. The vertebrate eye generally contains two kinds of photoreceptors which, because of their respective shapes, are known as rods and cones. In man, this duality of structure is paralleled physiologically by the presence of two, qualitatively different, kinds of vision. Visual sensation is entirely achromatic in dim lights while in brighter illumination vision is accompanied by color perception. Following the suggestion made by Parinaud[16] in 1881, v. Kries (1894)[11] has brought together much evidence to show that these two types of vision are each associated with a different kind of receptor. His Duplicity Theory, which is now universally accepted, supposes that in all vertebrates the rods alone are functional in dim vision and that the cones are associated with perception at higher intensities.

Since the eyes of fish have both rods and cones (Wunder, 1925),[18] it is very likely, therefore, that different mechanisms are concerned in the vision of fish at dim and bright illuminations. Further proof is given by the fact that color perception, which has been demonstrated in fish by v. Frisch (1913),[4] Mast (1915)[13] and others, is only possible at intensities

well above the threshold for brightness perception (v. Frisch, 1924).[5] Bauer (1910)[1] has also shown that fish exhibit the Purkinje phenomenon. Since this effect occurs in man because the rods and cones have different spectral sensibilities (Koenig and Ritter, 1891)[9] a similar condition probably exists in fish and it is possible to obtain direct evidence of the duplex nature of the fish visual system by showing that fish have different visibility functions at high and at low intensities. Furthermore, the nature of the photochemical substance concerned in bright vision may be elucidated by accurate knowledge of this function.

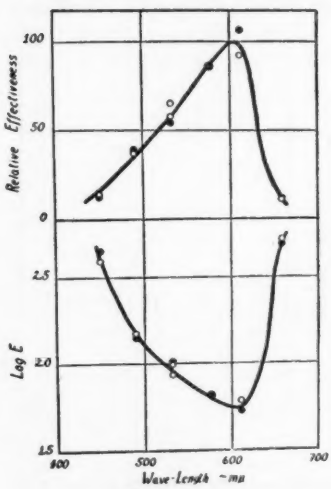


FIGURE 4

Cone visibility function for animal 6NS determined at two brightness levels. The solid circles are for a visual acuity of 0.06; the clear circles for a visual acuity of 0.12. In the latter case the points have been lowered 0.3 units on the log E ordinates to permit superposition of the two sets at 570 mμ.

With these aims in view, the visibility curve of Lepomis for bright illuminations has therefore been determined. It has been obtained by using a screen made up of very fine wires in the apparatus which was described above. The minimum intensities at which the animal responds to a movement of this screen are about 100–200 times higher than in the experiments with dim illuminations. Such bright spectral lights cannot be gotten with the spectrometer and instead spectral bands were obtained by using a series of Wratten "monochromat" filters in conjunction with a 500-watt Mazda lamp. The energy of the various bands has been computed from the energy emitted by the lamp and the spectral transmission of the filters. The intensity is again controlled by a neutral wedge and the experimental procedure consists in the determination of the wedge setting for each filter at which a response is just obtained.

5. The visibility curves of 4 animals have been measured in these experiments. Figure 4 shows the results obtained on one individual. The curve has a maximum near 600 mμ, showing that the spectral sensitivity for bright lights is shifted toward the red end of the spectrum in comparison with the sensibility in dim vision. The shift is in the same direction as that which occurs in man under similar conditions and shows that Lepomis has different spectral sensibilities at dim and at bright illuminations. It can therefore be concluded that

there are separate visual mechanisms concerned in vision at these two levels.

The curve shown in figure 4 has been obtained by using two different screens. The filled circles represent data for a screen made of wire having a diameter of 0.2 mm., while the clear circles are measurements with a screen in which the wire is 0.1 mm. in diameter. The visual acuities for these screens are 0.06 and 0.12 Snellen Units, respectively. The shape of the visibility curve obtained with these screens is the same although the brightness levels are different. This indicates that these measurements represent the visibility function of the mechanism responsible for bright vision and not an intermediate stage where the spectral characteristics of both the rods and cones are involved.

6. One further point is of interest. What, if any, relation is there between the photochemical substances which must exist in these two systems?

Reasons have been given above for supposing that there are present in the eye of Lepomis several pigments that are not light sensitive. Because these pigments absorb part of the light which enters the eye, the visibility curves do not actually correspond with the absorption curves of the photochemical substances involved in vision, so that no definite conclusions can be drawn. However, the bright visibility curve of one animal is much broader than the others that have been determined and is very similar to the exceptional curve obtained in the experiments at dim illuminations. Indeed, in figure 5, the clear circles represent the data for No. 3 at high illuminations and the filled circles those for No. 8I, for dim lights, shifted 30 mμ toward the red. The curve for No. 8I, which is also shifted to this extent describes approximately both sets of data. It has been proposed above that animal No. 8 had lost temporarily one of the pigments so that the violet portion of its scotopic visibility curve approaches the visual purple absorption curve. From the fit of this curve to these measurements on another animal at high intensities, it seems likely that the photochemical substances concerned in both bright and dim vision have similar absorption curves, but that the absorption curves of the substances mediating bright vision lie further toward the red end of the spectrum.

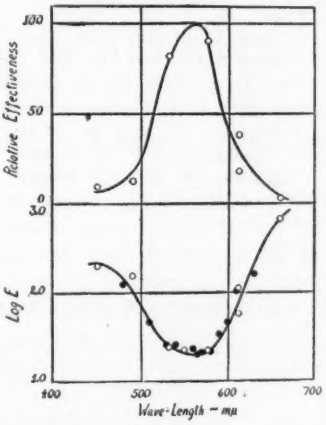


FIGURE 5

Cone visibility (clear circles) of animal 3NS. The solid circles are the data for animal 8 (Fig. 3, lower half) shifted 30 mμ toward the red.

* A part of this work was accomplished during tenure of a Fellowship in Biology from the NATIONAL RESEARCH COUNCIL.

[1] Bauer, *Pflug. Arch.*, **133**, 1910; **137**, 1911.
[2] Cunningham and MacMunn, *Phil. Trans. Roy. Soc. Lond.*, **B184**, 1893.
[3] Dieter, W., *Pflug. Arch.*, **222**, 1929.
[4] Frisch, K. v., *Zoöl. Jahrb. Abt. Physiol.*, **34**, 1913.
[5] Frisch, K. v., *Verh. dtsch. zoöl. Ges.*, **29**, 1924; *Z. vergl. Physiol.*, **2**, 1925.
[6] Hecht, S., and Williams, R. E., *Jour. Gen. Physiol.*, **5**, 1922.
[7] Hecht, S., and Wolf, *Ibid.*, **12**, 1929.
[8] Koenig, A., *Sitzungsber. Akad. Wiss. Berl.*, 1894, reprinted in *Gesamm. Abhandl.*, Leipzig, 1903.
[9] Koenig, A., und Ritter, R., *Z. Psychol. Physiol. Sinnesorg.*, **7**, 1891.
[10] Koettgen, E., und Abelsdorff, G., *Ibid.*, **12**, 1896.
[11] Kries, J. v., *Ber. d. Freiburg. Ges.*, **9**, 1894. General article in *Nagel's Handb. d. Physiol.*, **3**.
[12] Lyon, E. P., *Am. Jour. Physiol.*, **4**, 1904.
[13] Mast, S. O., *Proc. Nat. Acad. Sci.*, **1**, 1915; *Bull. U. S. Bureau of Fish.*, **35**, 1916.
[14] Murr, E., *Zoöl. Anz.* **3**, *Supplbd.*, 1928.
[15] Palmer, L. S., *Carotinoids and Related Pigments*, N. Y., 1922.
[16] Parinaud, H., *Arch. gen de Med.*, April, 1881.
[17] Trendelenburg, W., *Z. Psychol. Physiol. Sinnesorg.*, **36**, 1905.
[18] Wunder, W., *Z. Vergl. Physiol.*, **2**, 1925; **3**, 1926.
[19] Verne, J., *Les Pigments dans l'Organisme animal*, Paris, 1926.

SYSTEMS OF ALGEBRAIC DIFFERENTIAL EQUATIONS

BY J. F. RITT

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated April 29, 1931

In a recent paper,[1] we introduced the notion of irreducible system of algebraic differential equations and showed that every system of algebraic differential equations is equivalent to a finite number of irreducible systems. We showed also that the solution of an irreducible system could be made to depend on the solution of a single differential equation, called the *resolvent* of the system.

Our proofs were all of a purely abstract nature, and did not lead to methods for decomposing a system into irreducible systems or for constructing resolvents. To sketch such methods is the object of the present note.

We use the notation of our above-mentioned paper, which paper we shall designate by M. Consider a system of algebraic differential equations in the indeterminates $u_1, \ldots, u_q; y_1, \ldots, y_p$

$$\alpha_1 = 0, \ \ldots, \ \alpha_p = 0, \tag{1}$$

where

(a) Each α_i is a form in $u_1, \ldots, u_q; y_1, \ldots, y_i$, effectively involving y_i, with coefficients which are meromorphic functions belonging to a field $\mathfrak{F}$. For $j < i$, the order of α_i in y_j does not exceed that of α_j.

(b) If α_i is of order r_i in y_i, the coefficient of the highest power of y_{i,r_i} in α_i does not vanish when each y_{j,r_j}, $j < i$, is replaced by its expression in terms of the quantities

$$x;\ y_1, y_{11}, \ldots, y_{1,r_1-1};\ \ldots;\ y_j, y_{j1}, \ldots, y_{j,r_j-1};\ u_1, \ldots, u_q;\ u_{11}, \text{ etc.} \quad (2)$$

obtained from the equations $\alpha_k = 0$, $k = 1, \ldots, j$, considered as algebraic equations for $y_{1,r_1}, \ldots, y_{j,r_j}$.

(c) In the field obtained by adjoining to $\mathfrak{F}$ the quantities

$$y_j, \ldots, y_{j,r_j-1}, j = 1, \ldots, i;\ u_1, \ldots, u_q;\ u_{11}, \text{ etc.} \quad (3)$$

and also the $i - 1$ functions y_{j,r_j}, $j = 1, \ldots, i - 1$, described in (b), α_i is an irreducible polynomial in y_{i,r_i}, $i = 1, \ldots, p$.

Let $F_i = \partial\alpha_i/\partial y_{i,r_i}$. It can be proved that *the totality of those forms which vanish for all solutions of* (1) *for which no* F_i *vanishes constitute an irreducible system.*

Suppose now that we are given any finite system Σ of forms in the indeterminates $y_1, \ldots, y_n$, some of the forms being distinct from zero. If Σ contains a non-zero form of class zero, any such form in Σ will be called a *primary ascending set* of Σ. Suppose that Σ contains no such form. Let α_1, of class $p_1 > 0$, be one of those non-zero forms of Σ whose rank is a minimum. Suppose that Σ contains forms of class greater than p_1 and of lower rank than α_1 in y_{p_1}. Of all such forms, let α_2 have a minimum rank. Continuing, we obtain a system

$$\alpha_1, \ldots, \alpha_s \quad (4)$$

where

(a) Each α_i is a form of Σ of class $p_i > 0$.

(b) α_1 has a rank not greater than that of any other non-zero form of Σ.

(c) If $s > 1$, α_j with $j > 1$ is of higher class than α_i, $i = 1, \ldots, j - 1$, is of lower rank than every such α_i in y_{p_i}, and is of as low a rank as it can be with these assumptions.

(d) Σ has no form of class greater than p_s and of lower rank than each α_i in y_{p_i}.

The set (4) will be called a *primary ascending set* of Σ.

If we reduce the forms of Σ with respect to the forms of (4) by the method frequently used in M, we get a system of remainder forms, each of which holds Σ. Let Σ' be the system obtained by adjoining these remainder forms to Σ. Then Σ and Σ' hold each other. Now if some of the remainder forms are not zero, Σ' will have primary ascending sets which, in a sense easy to describe, are of lower rank than those of Σ.

Using this idea, we convert Σ, by a finite number of rational operations and differentiations, into a system Ω, with a primary ascending set reduction with respect to which yields only zero remainders. With suitable changes of notation, the primary ascending set for Ω can be considered as a system (1) for which condition (a) is satisfied. If (b) and (c) are not both satisfied, a process of factorization decomposes Ω into several systems with primary ascending sets of lower rank than those of Ω. We arrive thus, in a finite number of steps, at a finite set of systems, $\Sigma_1, \ldots, \Sigma_l$, with primary ascending sets of the type (1) for which (a), (b), (c) are satisfied, such that Σ holds each Σ_i and every solution of Σ is a solution of some Σ_i. The content of Σ_i will be the content of the irreducible system mentioned in connection with (1), together with the contents of the systems $\Sigma_i + F_j, j = 1, \ldots, p$. Now each of the latter systems has primary ascending sets of lower rank than those of Σ.

By successive reductions of the above types, we arrive at a set of systems (1) which are the basic systems of a finite number of irreducible systems equivalent to Σ. Our algorithm does not identify the irreducible systems completely, but gives only their basic systems. However, this is sufficient for the construction of the resolvents of the irreducible systems, as one will see now immediately on examining §§ 21–23 of *M*.

[1] Manifolds of functions defined by systems of algebraic differential equations. *Trans. Amer. Math. Soc.*, **32**, 569 (1930).

THE NUMBER OF REPRESENTATIONS FUNCTION FOR POSITIVE BINARY QUADRATIC FORMS

By Gordon Pall[1]

Department of Mathematics, California Institute of Technology

Communicated April 30, 1931

Let $f_1, \ldots, f_h$ be a representative set of positive, primitive, integral, binary, quadratic forms of a given discriminant $d < -4$. For any form f let $f(n)$ denote what has generally been designated by $N(n = f)$, namely, the number of representations of n by f. As long ago as 1840, Dirichlet found the value of the function

$$r(n) \equiv {}^1/_2\{f_1(n) + f_2(n) + \ldots + f_h(n)\}, \tag{1}$$

when n is prime to d, as the following simple function of the divisors of n:

$$r(n) = \Sigma(d/\nu) \tag{2}$$

summed for the positive divisors ν of n.[2]

Only when there is just one class in each genus of primitive forms does Dirichlet's theorem furnish the number $f_i(n)$ of representations in the individual forms f_i. This case occurs for one hundred and two known discriminants and no further ones up to $-23{,}000$. Till now nothing seems to have been published on the value of $f_i(n)$ in general.

Recently the writer has reduced the problem of evaluating the functions $f_i(n)$ to that of knowing, first, the structure of the primitive order with respect to composition, second, in which of the forms f_j the various primes dividing n are represented. Methods have been developed of finding simple closed formulae for the number of representations in *any integral, positive, binary, quadratic form whatever.*

We shall now state the final formula, when n is prime to d, in the simplest case: $h = 3$. Then there can be only one genus and its classes may be labeled G^0, the principal class, and G^1 and G^2, two classes mutually inverse under composition. For example, consider $d = -44$, when

$$f_1 = x^2 + 11y^2,\ f_2 = 3x^2 + 2xy + 4y^2,\ f_3 = 3x^2 - 2xy + 4y^2. \quad (3)$$

Generally, let f_1 denote the principal form, and f_2, f_3 representatives of the classes G^1, G^2, respectively. Evidently $f_2(n) = f_3(n)$ for every n. To find $f_i(n)$ write

$$n = p_1^{a_1} p_2^{a_2} \ldots p_t^{a_t} n_0 \quad (4)$$

where the p_j are all the distinct primes dividing n which are represented in f_2. For $j = 1, \ldots, t$ write

$$a_j = 3\alpha_j + k_j - 1,\ k_j = -1, 0, \text{ or } 1. \quad (5)$$

Then, n being prime to d,

$$f_1(n) = {}^2/_3 r(n) \left\{ 1 + \frac{2k_1k_2 \ldots k_t}{(a_1 + 1) \ldots (a_t + 1)} \right\}, \quad (6)$$

and

$$f_2(n) = f_3(n) = {}^2/_3 r(n) \left\{ 1 - \frac{k_1k_2 \ldots k_t}{(a_1 + 1) \ldots (a_t + 1)} \right\}. \quad (7)$$

In particular, if $t = 0$, $f_1(n) = 2r(n)$, $f_2(n) = f_3(n) = 0$; and if n is divisible by a single prime p_j (for example, by 3 in case (3)) to an exponent of the form $3\alpha - 1$, then $f_1(n) = f_2(n) = f_3(n) = {}^2/_3 r(n)$. Beautiful relations of similar types appear in general. Also, the formulae for $f_i(n)$ are easily extensible to the case where n is not necessarily prime to d.

These results arose from a much less ambitious investigation, the evaluation of $r(n)$ for n not necessarily prime to d. This was facilitated by showing first that $r(n)$ is factorable. Then reduction formulae of the type

$$f(pn) + f(n/p) = fg(n) + fg^{-1}(n) \quad (p \text{ in } g)$$

were found, and from them

$$f(p^a m) = fg^a(m) + fg^{a-2}(m) + \ldots + fg^{-a}(m).$$

These form the foundation of the theory.

The writer wishes to acknowledge his debt to Professor E. T. Bell without whose encouragement and suggestions these considerations might never have been carried to their logical conclusion.

[1] NATIONAL RESEARCH FELLOW.

[2] See, for example, Theorem 64, p. 78, of L. E. Dickson's *Introduction to the Theory of Numbers*.

DYNAMICAL TRAJECTORIES AND THE ∞^3 PLANE SECTIONS OF A SURFACE

BY EDWARD KASNER

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated April 5, 1931

This paper deals with a comparison of two fundamental types of triply infinite families of curves, the *dynamical* type and the *sectional* type. Both types are projectively invariant.

The *dynamical* type is obtained by taking an arbitrary positional field of force in the plane, so that the equations of motions are

$$\ddot{x} = \varphi(x, y), \ \ddot{y} = \psi(x, y) \qquad (1')$$

and considering the totality of possible trajectories for all initial conditions. The family of trajectories is defined by a differential equation of the particular form[1]

$$(\psi - y'\varphi)y''' = \{\psi_x + (\psi_y - \varphi_x)y' - \varphi_y y'^2\}y'' - 3\varphi y''^2. \qquad (1)$$

The *sectional* type is obtained, on the other hand, by taking on an arbitrary surface $z = f(x, y)$, all the plane sections $z = ax + by + c$, and projecting these sections orthogonally on the xy plane. The resulting family of curves is

$$ax + by + c - f(x, y) = 0. \qquad (2')$$

The differential equation of this family is easily found to be

$$(f_{xx} + 2f_{xy}y' + f_{yy}y'^2)y''' = (f_{xxx} + 3f_{xxy}y' + 3f_{xyy}y'^2 + f_{yyy}y'^3)y'' + 3(f_{xy} + f_{yy}y')y''^2. \qquad (2)$$

It is observed that both types are of the general form

$$y''' = G(x, y, y')y'' + H(x, y, y')y''^2 \qquad (3)$$

which is characterized by the simple property I stated later. (I may observe that I have met this form (3) in several other geometric and physical connections which will be discussed in a later paper. Examples are curvature trajectories and general catenaries.)

While the types (1) and (2) are of the same general form, they are of course distinct, as we see by observing the different manner in which y' is involved.

An example of a sectional family which is not dynamical is given by the ∞^3 circles of the plane. These are obtained from the paraboloid of revolution $z = x^2 + y^2$, but cannot be obtained from any field of force.

An example of a dynamical family which is not sectional is

$$Ax^2 + Bxy + Cy^2 = 1,$$

that is, all conics with center at the origin. This is the trajectory family of the central force varying directly as the distance, but cannot be obtained from any surface.

An example of a family belonging to both types is given by taking all conics with a focus at the origin. This is, of course, the trajectory family of the Newtonian field; and can be identified with the orthogonal projections of the plane section of the cone of revolution.

Our main problem[2] is to find *all families of curves which are of both dynamical and sectional types.* We find all surfaces whose plane sections give rise by orthogonal projection to a trajectory family; and all fields of force whose trajectories can be related to a surface.

We show that *the only surfaces which give rise to trajectories are the general cone and cylinder,* and that *the corresponding fields of force are central and parallel, respectively, of the particular types given by formulas* (11*a*) *and* (11*b*) below. The circular cone produces the trajectories of the Newtonian inverse square field, and the parabolic cylinder those of the parallel gravitational field. *No other conservative field can be generated by a surface.*

The restriction to orthogonal projection is then removed, and we prove finally that *the projections from any center on to any plane of the plane sections of a surface form a system of dynamical trajectories if and only if the surface is a cone* (*of any shape*).

We observe that all the concepts involved in the paper are projectively invariant, and therefore the results contribute new material to the *projective differential geometry of surfaces and of triply infinite families of plane curves.* (The sectional type is also of topological interest as noted at the end of this paper.)

Determination of All Surfaces That Generate Trajectories.—1. It was proved in the Transaction paper referred to that *a system of dynamical trajectories may be characterized by five geometrical properties.*[1] It is convenient to use six properties, I to VI, but III and IV are equivalent.

PROPERTY I: If to each of the ∞^1 curves having a given lineal element in common the osculating parabola is drawn at that element, the foci will lie on a circle through the point of the element. This property is characteristic for equations of the form (3).

PROPERTY II: There exists for each point (x, y) of the plane a certain direction ω (that of the "*force*") such that the angle between this direction and the tangent to the focal circle corresponding to any element (x, y, y') at the given point, is bisected by that element. This is equivalent to requiring the function $H(x, y, y')$ in (3) to have the form $\frac{3}{y' - \omega(x, y)}$.

PROPERTY III: The locus of the centers of the ∞^1 circles corresponding to the lineal elements at a given point is a conic with that point as a focus. Properties I, II and III are possessed only by systems of curves whose differential equation is of the form

$$(y' - \omega)y''' = (\lambda y'^2 + \mu y' + \nu)y'' + 3y''^2 \qquad (4)$$

where ω, λ, μ, ν are arbitrary functions of x, y.

PROPERTY IV: In each direction through a given point O there passes one trajectory which has contact of third order with its circle of curvature. The locus of the centers of the ∞^1 hyperosculating circles, obtained by varying the initial direction, is a conic passing through the given point in the direction of the force. This property is completely equivalent to property III, so that type (4) has properties I, II, III, IV.

PROPERTY V: Of the curves which pass through a given point in the direction of the "force" at that point, there is one which has contact of the third order with its circle of curvature; the radius of curvature of this curve is three times the radius of curvature of the "line of force" (i.e., integral curve of the direction assigned to each point by II) passing through the given point. This property in analytic form requires the functions in (4) to obey the relation

$$\lambda\omega^2 + \mu\omega + \nu + \omega_x + \omega\omega_y = 0. \qquad (5)$$

PROPERTY VI: When the point O is moved, the associated conic described in property IV changes in the following manner. Take any two fixed perpendicular directions for the x direction and the y direction; through O draw lines in these directions meeting the conic again at A and B, respectively. Also construct the normal at O meeting the conic again at N. At A draw a line in the y direction meeting this normal in some point A', and at B draw a line in the x direction meeting the normal in some point B'. The variation property referred to, takes the form

$$\frac{\partial}{\partial x}\frac{1}{AA'} + \frac{\partial}{\partial y}\frac{1}{BB'} + \frac{\omega\omega_{xy} - \omega_x\omega_y}{3\omega^2} = 0,$$

where AA' and BB' denote distances between points, and where ω denotes the slope of the lines of force relative to the chosen x direction. This is true for any pair of orthogonal directions, and therefore, really expresses an intrinsic property of the system of curves. (See the diagram in references.[1])

The analytic form of this property, which we shall use later, is

$$\lambda_x + \left(\frac{\nu + \omega_x}{\omega}\right)_y = 0 \tag{6}$$

where subscripts as throughout the paper denote partial derivatives.

2. We apply these five conditions to restrict the surface $z = f(x, y)$ so that (2) may be of the form (1).

Property I is satisfied by (2) always, since (2) is a special case of (3). We therefore have

THEOREM I. *The curves produced by orthogonal projection of the plane sections of any surface have Property I. That is, for any lineal element the focal locus is a circle through the point of the element.*

Property II requires $\omega(x, y) = -\dfrac{f_{xx} + f_{xy}y'}{f_{xy} + f_{yy}y'}$ to be independent of y'. Therefore

$$f_{xx}f_{yy} - f_{xy}^2 = 0, \tag{7}$$

and therefore *the surface must be developable.*

Property III (with IV) is obeyed by (2) in virtue of (7) and

$$\begin{aligned} f_{yy}f_{xxx} + f_{xx}f_{xyy} - 2f_{xy}f_{xxy} &= 0 \\ f_{yy}f_{xxy} + f_{xx}f_{yyy} - 2f_{xy}f_{xyy} &= 0 \end{aligned} \tag{8}$$

obtained by differentiating (7). We find that the coefficients of y''' and y'' in (2) have the common factor $f_{xy} + f_{yy}y'$ and that

$$\begin{gathered} \omega = -\frac{f_{xx}}{f_{xy}}, \quad \lambda = \frac{f_{yyy}}{f_{yy}} \\ \mu = \frac{3f_{xyy}f_{yy} - f_{xy}f_{yyy}}{f_{yy}^2}, \quad \upsilon = \frac{f_{xxx}}{f_{xy}}. \end{gathered} \tag{9}$$

Substituting in (5), and applying (7), we find

$$f_{xxx}f_{yy}^2 - f_{xxy}f_{xy}f_{yy} - f_{xyy}f_{xx}f_{yy} + f_{yyy}f_{xx}f_{xy} = 0$$

which relations (8) show to hold. This verifies property V, and establishes

THEOREM II. *The curves produced by orthogonal projection of the plane sections of any developable surface have properties I, II, III, IV and V.*

Property VI, in the form (6), imposes on f the extra condition of fourth order

$$\left(\frac{f_{yyy}}{f_{yy}}\right)_x - \left(\frac{f_{xxy}}{f_{xy}}\right)_y = 0$$

as we find by means of (7) and (8). We find, after a rather complicated discussion requiring integrations, that the surface is either a cone or a cylinder. We may reduce the equation of the surface in fact to one of the two forms

$$z = xh\left(\frac{y}{x}\right) \tag{10a}$$

$$z = g(x) \tag{10b}$$

where h and g are arbitrary functions. By comparison with (1), we deduce that the fields of force corresponding to these surfaces are defined by

$$\varphi = x^{-2}h''(y/x),\ \psi = yx^{-3}h''(y/x) \tag{11a}$$

$$\varphi = 0,\ \psi = g''(x). \tag{11b}$$

Since $\psi/\varphi = y/x$ in (11a), a cone generates a central field. In polar coördinates the magnitude of the force is $r^{-2}\mathrm{h}^*(\theta)$ where h^* is an arbitrary function. That is, *along any direction from the center the force is of the inverse square type.* The cylinder (10a) produces a parallel field (11b) of constant intensity along any line of force.

THEOREM III. *The orthogonal projections of the plane sections of a general cone or cylinder form a set of curves which can be identified with the trajectories of a central or parallel field of force, respectively, of the special type described. No other surface can yield a dynamical family.*

3. If we impose the condition $\psi_x - \varphi_y = 0$ that the fields (11) are conservative, then we find that the only solutions are

$$\varphi = xr^{-3},\ \psi = yr^{-3}\ \text{(Newtonian field)} \tag{12a}$$

$$\varphi = 0,\ \psi = \text{constant (Constant field)}. \tag{12b}$$

The surfaces (10) then reduce, respectively, to the circular cone $z^2 = x^2 + y^2$ and the parabolic cylinder $z = x^2$.

THEOREM IV. *The orthogonal projections of the plane sections of a circular cone are the trajectories of the Newtonian field; those of a parabolic cylinder are the trajectories of a constant parallel field. No other surface generates the trajectories of a conservative field.*

Generalization to Central Projection.—It was shown in the reference[1] (p. 422) that the group of point transformations which convert every system of dynamical trajectories into such a system is the group of col-

lineations of the plane. This result enables us immediately to extend our result to any central projection of the plane sections of a surface, and to declare that the projected curves are dynamical trajectories when and only when the surface is a cone. For suppose the curves are trajectories. Let a projective transformation of space be made which carries the plane into some other plane, and the center of the projection into the point at infinity on the normal to the transformed plane. The transformed surface then generates trajectories by orthogonal projection, and so, from Theorem III, must be a cone (a term which may now include the cylinder). Hence, the original surface, which is derived by a projective transformation from a cone, must itself be a cone.

We showed in our Transaction paper[1] that the geometric properties I to VI are *separately* (as well as collectively) invariant under projective transformation. Moreover, a developable surface remains developable after such a transformation. Hence in theorems I and II we may understand central in place of orthogonal projection and the results are still valid.

THEOREM V. *Any projection of the plane sections of an arbitrary surface on a given plane yields a triply-infinite family of curves possessing property I (focal locus for any element is a circle passing through the point, or analytic form* (3)). *If the surface is developable, the family has properties I to V. If the surface is conical (of any cross section), the family has all the properties I to VI, and so we have a dynamical family.*

It has been shown above synthetically that the central projection of any surface S gives a triple family of curves which is projectively equivalent to the orthogonal projection family of some other surface S'. We can show, moreover, analytically, by applying a collineation to a family in its finite form $ax + by + c - f(x, y) = 0$, that the transformed family is of the same form, with merely a new function f. The differential equation is thus of the same form (2). Hence the family obtained by central projection from surface S can actually be obtained by orthogonal projection from some new surface S''. The surface S'' as well as S' is projectively equivalent to S.

THEOREM VI. *Our sectional type* (2), *as well as the dynamical type* (1), *is invariant under the collineation group of the plane.*

Hence we need make no distinction (when we take into account the totality of surfaces $z = f(x, y)$) between the ortho-sectional type and the centro-sectional type. In both cases we obtain differential equations of the same form (2), which we term simply the sectional type.

This type could be characterized geometrically by adjoining certain complicated projective properties[2] to the simple property I.

Any linear[3] three-parameter family of curves can, under the group of all point transformations, be reduced to the sectional form. Hence

there is an application to *topological* differential geometry which we shall consider elsewhere.

[1] E. Kasner, "The Trajectories of Dynamics," *Trans. Amer. Math. Soc.*, **7**, 401–424 (1906). Also "Differential-Geometric Aspects of Dynamics," Princeton Colloquium Lectures on Mathematics (1913), especially Chap. I and Chap. III, where some of the properties are stated in projective language.

[2] *Bull. Amer. Math. Soc.*, **14**, 356 (1908), where the results are stated. Also *Bull.*, **36**, 51 (1930). We assume that the field in (1) is not null, and that the surface in (2) is not a plane; these degenerate cases give merely the ∞^2 straight lines and so correspond.

[3] The term linear is here used in its general or topological, rather than its more usual projective or algebraic meaning. See author's St. Louis address, *Bull. Amer. Math. Soc.*, **11**, 307 (1905).

ON THE SUMMABILITY OF FOURIER SERIES. FOURTH NOTE

BY EINAR HILLE AND J. D. TAMARKIN

PRINCETON UNIVERSITY AND BROWN UNIVERSITY

Communicated April 23, 1931

1. In the present note we continue the discussion of the summability of Fourier series by the method $[H, q(u)]$ of Hurwitz-Silverman-Hausdorff.[1] We extend the discussion to the conjugate of a Fourier series and also to the derived of the conjugate of Fourier series of functions of bounded variation. A comparison between the results for Fourier series and for their conjugate series is rather striking.

We use the notation of our third note except for the following modifications. We put

$$\gamma(u) = q(u) - u, \quad 0 \leqq u \leqq 1, \tag{1}$$

$$C(\xi) = \int_0^1 \cos \xi u \gamma(u) du, \quad S(\xi) = \int_0^1 \sin \xi u \gamma(u) du, \tag{2}$$

$$J(\epsilon) = \int_\epsilon^1 \frac{Q(u)}{u}\, du + \int_0^{1-\epsilon} \frac{Q(1) - Q(u)}{1-u}\, du, \tag{3}$$

where $Q(u)$ is the total variation of $q(v)$ on $0 \leqq v \leqq u$.

The conjugate of the Fourier series of a function $f(x) \subset L$ is

$$\sum_{n=1}^{\infty} (-b_n \cos nx + a_n \sin nx), \tag{4}$$

where

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx). \tag{5}$$

On setting

$$c_n = a_n - ib_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt, \tag{6}$$

and denoting the partial sums of the series (4) and (5) respectively, by $\tilde{s}_n(x)$ and $s_n(x)$, we have

$$\sigma_n(x) = s_n(x) + i\tilde{s}_n(x) = \frac{c_0}{2} + \sum_{\nu=1}^{n} c_\nu e^{i\nu x}. \tag{7}$$

Assuming $f(x)$ to be real we obtain for the nth $[H, q]$ mean of the sequence $\{\sigma_n(x)\}$ the expression[2]

$$\mathfrak{H}_n[f(x), q] = \int_{-\pi}^{\pi} f(x+t)\mathfrak{H}_n(t)dt = H_n[f(x), q] + i\tilde{H}_n[f(x), q] \tag{8}$$

where

$$\mathfrak{H}_n(t) = \frac{i}{2\pi}\cot\frac{t}{2}\left[\int_0^1 (1-u+ue^{-it})^n dq(u) - 1\right]$$
$$+ \frac{1}{2\pi}\int_0^1 (1-u+ue^{-it})^n dq(u) = H_n(t) + i\tilde{H}_n(t). \tag{9}$$

A point x will be called $(\tilde{F})$ regular or $(\tilde{L})$ regular with respect to $f(x)$ if (i)x is a point of continuity of $f(x)$ or (i′)

$$\int_0^t |f(x+\tau) - f(x-\tau)|d\tau = o(t), \tag{10}$$

while, in both cases, (ii) the limit

$$\tilde{f}(x) \equiv -\frac{1}{2\pi}\lim_{\epsilon\to 0}\int_\epsilon^\pi [f(x+t) - f(x-t)]\cot\frac{t}{2}\,dt \tag{11}$$

exists and is finite. The points which were designated in the third note as regular or pseudo-regular will be called here (F) regular or (L) regular. It is well known that almost all points, and, in particular, all $(\tilde{F})$ regular points are $(\tilde{L})$ regular.

The nth $[H, q]$ mean of the conjugate series (4) is

$$\tilde{H}_n[f(x), q] = \int_0^\pi [f(x+t) - f(x-t)]\tilde{H}_n(t)\,dt. \tag{12}$$

The method $[H, q]$ will be said to be $(\tilde{F})$ effective or $(\tilde{L})$ effective if the series (4) is summable $[H, q]$ to the sum $\tilde{f}(x)$ for every $f(x) \subset L$ at every $(\tilde{F})$ regular or at every $(\tilde{L})$ regular point of $f(x)$, respectively.

The derived series require the following notation and terminology. We assume now $f(x)$ to be periodic and of bounded variation on $(-\pi, \pi)$. A point x will be called (L') regular with respect to $f(x)$ if (i) $f'(x)$ exists and (ii)[3]

$$\int_0^t |d\psi(\tau)| = o(t) \tag{13}$$

where

$$\psi(t) = f(x + t) - f(x - t) - 2t f'(x). \tag{14}$$

Similarly, a point x is defined to be $(\tilde{L}')$ regular with respect to $f(x)$ if (i) the limit

$$\frac{1}{4\pi} \lim_{\epsilon \to 0} \int_{\epsilon}^{\pi} \varphi(t) [\sin \frac{t}{2}]^{-2} dt \equiv \tilde{f}\,'(x) \tag{15}$$

exists and is finite and (ii)

$$\int_0^t |d\varphi(\tau)| = o(t) \tag{16}$$

where

$$\varphi_1(t) = f(x + t) + f(x - t) - 2 f(x). \tag{17}$$

It is well known[4] that, except for sets of measure zero, all points are (L') regular and $(\tilde{L}')$ regular.

The nth $[H, q]$ mean of the derived sequence $\{\sigma_n'(x)\}$ is

$$\mathfrak{H}_n'[f(x), q] = \int_{-\pi}^{\pi} \mathfrak{H}_n(t)\, d_t f(x + t) = H_n'[f(x), q] + i\tilde{H}_n'[f(x), q] \tag{18}$$

where $H_n[f(x), q]$ and $\tilde{H}_n[f(x), q]$ are the nth $[H, q]$ means of the series (5) and of its conjugate series (4), respectively. It is found that

$$H_n'[f(x), q] - f'(x) = \int_0^{\pi} H_n(t) d\psi(t), \tag{19}$$

$$H_n'[f(x), q] = \int_0^{\pi} \tilde{H}_n(t) d\varphi(t). \tag{20}$$

The method $[H, q]$ will be said to be (L') effective or $(\tilde{L}')$ effective if it sums the derived of the series (5) to the sum $f'(x)$ and the derived of the series (4) to the sum $\tilde{f}\,'(x)$ at all the points x which are (L') or $(\tilde{L}')$ regular, respectively, $f(x)$ being any periodic function of bounded variation on $(-\pi, \pi)$.

2. Our results for the conjugate series are stated in Theorem 1 below. For the sake of comparison we include the corresponding results for the ordinary Fourier series, of which a part was proved in our third note. The symbol $\subset L$ used in this section refers to the infinite interval $(-\infty, \infty)$. It is assumed that $q(u)$ satisfies the necessary and sufficient conditions of Hausdorff for the regularity of $[H, q]$. In addition we assume that

$$\lim_{\epsilon \to 0} J(\epsilon) \text{ is finite}, \tag{21}$$

where $J(\epsilon)$ is defined by (3).

THEOREM 1. *A sufficient condition for an* $[H, q]$ *satisfying* (21) *to be*

(*F*) *effective is that* $C(\xi) \subset L.$ (22.1)	($\tilde{F}$) *effective is that* $S(\xi) \supset L.$ (22.2)

A necessary condition for an $[H, q]$ *satisfying* (21) *to be*

(*L*) *effective is that* $\lim_{n\to\infty} n\int_{n-1}^{n-1/2} \varphi(t)\, C(nt)\, dt = 0,$ (23.1)	($\tilde{L}$) *effective is that* $\lim_{n\to\infty} n\int_{n-1}^{n-1/2} \psi_0(t) S(nt)\, dt = 0,$ (23.2)

which is equivalent to

$\lim_{n\to\infty} n^2\int_{n-1}^{n-1/2} \varphi_1(t) C'(nt)\, dt = 0,$ (24.1)	$\lim_{n\to\infty} n^2\int_{n-1}^{n-1/2} \psi_1(t) S'(nt)\, dt = 0,$ (24.2)

where

$$\psi_0(t) = f(x+t) \dot{-} f(x-t),\ \varphi_1(t) = \int_0^t \varphi(\tau)\, d\tau,\ \psi_1(t) = \int_0^t \psi_0(\tau)\, d\tau. \quad (25)$$

A sufficient condition for an $[H, q]$ *satisfying* (21) *to be*

(*L*) *effective is that* $\xi\lvert C(\xi)\rvert \leqq A(\xi),$ (26.1)	($\tilde{L}$) *effective is that* $\xi\lvert S(\xi)\rvert \leqq A(\xi),$ (26.2)

where $A(\xi)$ *is used as a generic notation to designate a function which is non-negative and of bounded variation on* $(1, \infty)$ *and such that the integral*

$$\int_1^\infty \frac{A(\xi)}{\xi}\, d\xi \quad (27)$$

is finite.

The results for the derived series are as follows:

THEOREM 2. *A necessary condition for an* $[H, q]$ *satisfying* (21) *to be*

(*L'*) *effective is that* $\lim_{n\to\infty} n\int_{n-1}^{n-1/2} C(nt)\, d\psi(t) = 0,$ (28.1)	($\tilde{L}'$) *effective is that* $\lim_{n\to\infty} n\int_{n-1}^{n-1/2} S(nt)\, d\varphi(t) = 0,$ (28.2)

which is equivalent to

$\lim_{n\to\infty} n^2\int_{n-1}^{n-1/2} C'(nt)\psi(t)\, dt = 0.$ (29.1)	$\lim_{n\to\infty} n^2\int_{n-1}^{n-1/2} S'(nt)\varphi(t)\, dt = 0.$ (29.2)

A sufficient condition for an $[H, q]$ *satisfying* (21) *to be* (*L'*) *or* ($\tilde{L}'$) *effective is* (26.1) *or* (26.2), *respectively, with the additional assumption that* $A(\xi)$ *be continuous on* $(1, \infty)$.

A necessary condition for a regular $[H, q]$ to be (*F*) effective is simply

$$C(\xi) \subset L, \quad (30.1)$$

as it has been proved in the third note. A natural question, whether

$$S(\xi) \subset L \tag{30.2}$$

is necessary for a regular $[H, q]$ to be $(\tilde{F})$ effective, remains open.

It would be of some interest also to exhibit a regular $[H, q]$ which is (F) effective without being $(\tilde{F})$ effective or vice-versa. We have reasons to believe that such a construction is possible.

3. Sufficient conditions for effectiveness of a different type are available on the basis of the discussion in our second note. We extend the definition of $q(u)$ outside of $(0, 1)$ by setting $q(u) = 0$ for $u < 0$ and $q(u) = 1$ for $u > 1$. We set

$$\Omega_Q(h) = \max_{0 \leq t \leq h} \int_{-\infty}^{\infty} |d_u[Q(u + t) - Q(u)]| \tag{31}$$

and assume that

$$\int_0^1 \Omega_Q(h) \frac{dh}{h} \text{ is finite.} \tag{32}$$

We have then

THEOREM 3. *Condition* (32) *is sufficient in order that a regular* $[H, q]$ *be* (F), $(\tilde{F})$, (L), $(\tilde{L})$, (L') *and* $(\tilde{L}')$ *effective.*

Condition (32) requires that $\Omega_Q(h) \longrightarrow 0$ as $h \longrightarrow 0$, which implies the absolute continuity of $q(u)$.[5] A particularly simple and important case is that in which $q(u)$ is a convex function satisfying (32). The corresponding methods $[H, q]$ which were mentioned in our second note appear to be equivalent, at least in so far as Fourier series are concerned, to a class of methods of summation studied by F. Nevanlinna from the point of view of (F) effectiveness.[6]

[1] Cf. these PROCEEDINGS **14**, 915–918 (1928); **15**, 41–42 (1929); and **16**, 594–598 (1930).

[2] $H_n(t)$ is the kernel of our earlier notes defined by the formulas (10) of the first and (2) of the third note. Both contain a misprint; the first exponental should be replaced by its conjugate. The third note contains a number of misprints. An annoying one occurs in formula (10) which should read $\rho(\xi)/\xi \subset L$. The statement of Theorem 4 is incomplete; it is necessary to assume that $g(u)$ vanishes outside of a finite interval or satisfies some similar additional condition at infinity.

[3] The notation here differs from that of the third note.

[4] A. Plessner, *Mitt. Math. Seminars Univ. Giessen*, Heft 10 (1923).

[5] A. Plessner, *J. Math.*, **160**, 26–32 (1929).

[6] *Översikt Finska Vet.-Soc. Förhandl.*, **64**, A, No. 3, 14 pp. (1921–1922).

ON THE STABILITY OF THE MOTION OF A VISCOUS FLUID

BY W. T. MACCREADIE

DEPARTMENT OF MATHEMATICS, BUCKNELL UNIVERSITY

Communicated April 24, 1931

A criterion for the stability of the steady motion of a viscous fluid between two parallel planes was established by Reynolds[1] as $2\rho b U_m/\mu < 517$, where ρ = density, $2b$ = distance between the parallel planes, U_m = mean velocity, and μ = coefficient of viscosity. A much lower figure 167 was found by Sharpe[2] and later Orr[3] obtained the figure 117. In this paper by expressing the small motion which is superimposed on the steady motion (the Reynold's method of solution) in the form of a Fourier's Series it is possible to show that the method employed by Reynolds gives a result as small as that of Orr. Also by applying Calculus of Variations the minimum value 116.8 is obtained.

Equations of Motion.—Consider a viscous fluid of density ρ moving in two dimensions with component velocities u, v, in directions x, y, respectively. The dynamical equations of motion are

$$\rho\frac{\partial u}{\partial t} = -\left\{\frac{\partial}{\partial x}(p_{xx} + \rho uu) + \frac{\partial}{\partial y}(p_{yx} + \rho uv)\right\} \qquad (1)$$

$$\rho\frac{\partial v}{\partial t} = -\left\{\frac{\partial}{\partial x}(p_{xy} + \rho vu) + \frac{\partial}{\partial y}(p_{yy} + \rho vv)\right\},$$

where the stresses in the fluid are given by

$$p_{xx} = p - 2\mu\frac{\partial u}{\partial x}, \quad p_{yy} = p - 2\mu\frac{\partial v}{\partial y}, \quad p_{xy} = p_{yx} = \mu\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right). \qquad (1a)$$

We have also the equation of continuity $\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} = 0$.

Steady Motion.—In the case of a fluid moving between two parallel planes there exists a solution of these equations of the form

$$v = 0, \; u = \frac{1}{\mu}\frac{\partial p}{\partial x}\frac{y^2 - b^2}{2}, \qquad (2)$$

where $y = \pm b$ are the equations of the planes and $\dfrac{\partial p}{\partial x}$ is constant. The equations (1) for this case reduce to

$$\frac{\partial p}{\partial x} = \mu\frac{d^2u}{dy^2}. \qquad (2a)$$

The average velocity is

$$U_m = \frac{\int_a^b u dy}{b} = -\frac{1}{3\mu}\frac{\partial p}{\partial x} b^2,$$

therefore $\frac{\partial p}{\partial x} = -3\mu U_m/b^2, \quad u = \frac{3}{2} U_m(b^2 - y^2)/b^2.$

Turbulent Motion.—Consider a motion given by $u = \bar{u} + u'$ and $v = \bar{v} + v'$ where u, v are periodic functions of period l with regards to x and also functions of y such that

$$\int_0^l u'dx = \int_0^l v'dx = 0 \text{ and } \bar{u} = \int_0^l u dx, \bar{v} = \int_0^l v dx.$$

To obtain the equations of turbulent motion substitute for u, v, in equations (1) and take the mean value of each member over a length l in the x-direction. The resulting equations for the mean motion are

$$\begin{aligned} \rho\frac{\partial \bar{u}}{\partial t} &= -\left\{\frac{\partial}{\partial x}(\bar{p}_{xx} + \rho\bar{u}\bar{u} + \rho\overline{u'u'}) + \frac{\partial}{\partial y}(\bar{p}_{yx} + \rho\bar{u}\bar{v} + \rho\overline{u'u'})\right\} \\ \rho\frac{\partial \bar{v}}{\partial t} &= -\left\{\frac{\partial}{\partial x}(\bar{p}_{xy} + \rho\bar{v}\bar{u} + \rho\overline{v'u'}) + \frac{\partial}{\partial y}(\bar{p}_{yy} + \rho\bar{v}\bar{v} + \rho\overline{v'v'})\right\}, \end{aligned} \tag{3}$$

where $p_{ij} = \bar{p}_{ij} + p'_{ij}$. Subtracting equations (3) from (1) we have the equations of relative motion,

$$\begin{aligned} \rho\frac{\partial u'}{\partial t} &= -\left\{\frac{\partial}{\partial x}[p'_{xx} + \rho(\bar{u}u' + u'\bar{u}) + \rho(u'u' - \overline{u'u'})] + \right. \\ &\qquad \left. \frac{\partial}{\partial y}[p'_{yx} + \rho(\bar{u}v' + u'\bar{v}) + \rho(u'v' - \overline{u'v'})]\right\} \\ \rho\frac{\partial v'}{\partial t} &= -\left\{\frac{\partial}{\partial x}[p'_{xy} + \rho(\bar{v}u' + v'\bar{u}) + \rho(v'u' - \overline{v'u'})] + \right. \\ &\qquad \left. \frac{\partial}{\partial y}[p'_{yy} + \rho(\bar{v}v' + v'\bar{v}) + \rho(v'v' - \overline{v'v'})]\right\}. \end{aligned} \tag{4}$$

To obtain the rate of change of the energy for the mean motion multiply equation (3) by $\bar{u}, \bar{v}$ and add. Putting $2\bar{E} = \rho(\bar{u}^2 + \bar{v}^2)$ we obtain

$$\frac{\partial \bar{E}}{\partial t} = -\left\{\begin{matrix} \frac{\partial}{\partial x}[\bar{u}(\bar{p}_{xx} + \rho\overline{u'u'})] + \frac{\partial}{\partial y}[\bar{u}(\bar{p}_{yx} + \rho\overline{u'v'})] \\ \frac{\partial}{\partial x}[\bar{v}(\bar{p}_{xy} + \rho\overline{v'u'})] + \frac{\partial}{\partial y}[\bar{v}(\bar{p}_{yy} + \rho\overline{v'v'})] \end{matrix}\right\} +$$

$$\begin{Bmatrix} \bar{p}_{xx}\dfrac{\partial \bar{u}}{\partial x} + \bar{p}_{yx}\dfrac{\partial \bar{u}}{\partial y} \\ \bar{p}_{xy}\dfrac{\partial \bar{v}}{\partial x} + \bar{p}_{xy}\dfrac{\partial \bar{v}}{\partial y} \end{Bmatrix} + \rho \begin{Bmatrix} \overline{u'u'}\dfrac{\partial \bar{u}}{\partial x} + \overline{u'v'}\dfrac{\partial \bar{u}}{\partial y} \\ \overline{v'u'}\dfrac{\partial \bar{v}}{\partial x} + \overline{v'v'}\dfrac{\partial \bar{v}}{\partial y} \end{Bmatrix}. \tag{5}$$

In a similar manner to obtain the equation for the rate of change of mean energy of the relative motion multiply (4) by u', v', and add. Omitting terms whose average value is zero and putting $2\bar{E}' = \rho(\overline{u'u'} + \overline{v'v'})$ we have

$$\frac{\partial \bar{E}'}{\partial t} = -\begin{Bmatrix} \dfrac{\partial}{\partial x}[u'(\bar{p}'_{xx} + \rho u'u') + \dfrac{\partial}{\partial y}[u'(\bar{p}'_{yx} + \rho u'v'] \\ \dfrac{\partial}{\partial x}[v'(\bar{p}'_{xy} + \rho v'u') + \dfrac{\partial}{\partial y}[v'(\bar{p}_{yy} + \rho v'v'] \end{Bmatrix} + \begin{Bmatrix} \bar{p}'_{xx}\dfrac{\partial u'}{\partial x} + \bar{p}_{yx}\dfrac{\partial u'}{\partial y} \\ \bar{p}'_{xy}\dfrac{\partial v'}{\partial x} + \bar{p}_{yy}\dfrac{\partial v'}{\partial y} \end{Bmatrix} - \rho \begin{Bmatrix} \overline{u'u'}\dfrac{\partial \bar{u}}{\partial x} + \overline{u'v'}\dfrac{\partial \bar{u}}{\partial y} \\ \overline{v'u'}\dfrac{\partial \bar{v}}{\partial x} + \overline{v'v'}\dfrac{\partial \bar{v}}{\partial y} \end{Bmatrix}. \tag{6}$$

If we integrate (5) and (6) between the planes, the last terms in each gives the conversion of energy of mean motion into energy of relative motion and is the same except for sign. After substituting the value from (1a) we have from (6)

$$\begin{aligned} \int_{-b}^{b}\int_{0}^{l} \frac{d\bar{E}'}{dt}\,dxdy = &-\rho \int_{-b}^{b}\int_{0}^{l} \begin{Bmatrix} \overline{u'u'}\dfrac{\partial \bar{u}}{\partial x} + \overline{u'v'}\dfrac{\partial \bar{u}}{\partial y} \\ \overline{v'u'}\dfrac{\partial \bar{v}}{\partial x} + \overline{v'v'}\dfrac{\partial \bar{v}}{\partial y} \end{Bmatrix} dxdy \\ &-\mu \int_{-b}^{b}\int_{0}^{l} \begin{Bmatrix} 2\left[\left(\dfrac{\partial u'}{\partial x}\right)^2 + \left(\dfrac{\partial v'}{\partial y}\right)^2\right] \\ +\left(\dfrac{\partial v'}{\partial x} + \dfrac{\partial u'}{\partial y}\right)^2 \end{Bmatrix} dxdy. \end{aligned} \tag{7}$$

In order that the relative motion may not die out it is necessary that the right-hand member of (7) be greater than or equal to zero, or

$$\int_{-b}^{b}\int_{0}^{l} \frac{d\bar{E}'}{dt}\,dxdy = \rho I_1 - \mu I_2 \geq 0, \tag{8}$$

where I_1, I_2, are the double integrals in equation (7). Hence, values of u', v' must exist for which I_2/I_1, must be $\geq \rho/\mu$ for a minimum value. They must also satisfy the continuity equations $\partial u'/\partial x + \partial v'/\partial y = 0$ and the boundary conditions $u' = v' = 0$ when $y = \pm b$.

Assuming
$$\begin{cases} u' = \dfrac{d\alpha}{dy}\cos\dfrac{2\pi x}{l} + \dfrac{d\beta}{dy}\sin\dfrac{2\pi x}{l} \\ v' = \dfrac{2\pi}{l}\alpha\sin\dfrac{2\pi x}{l} - \dfrac{2\pi}{l}\beta\cos\dfrac{2\pi x}{l} \end{cases}$$

where $\alpha = \beta = \dfrac{d\alpha}{dy} = \dfrac{d\beta}{dy} = 0$, when $y = \pm b$, substituting for u' and v' in I_1, I_2 and integrating over the length l in the x-direction we have

$$I_1 = \frac{l}{2}\int_{-b}^{b}\left\{\frac{2\pi}{l}\left(\alpha\frac{d\beta}{dy} - \beta\frac{d\alpha}{dy}\right)\frac{d\bar{u}}{dy}\right\}dy, \tag{8a}$$

$$I_2 = \frac{l}{2}\int_{-b}^{b}\left\{\left(\frac{2\pi}{l}\right)^4(\alpha^2+\beta^2) + 2\left(\frac{2\pi}{l}\right)^2\left[\left(\frac{d\alpha}{dy}\right)^2 + \left(\frac{d\beta}{dy}\right)^2\right] + \left(\frac{d^2\alpha}{dy^2}\right)^2 + \left(\frac{d^2\beta}{dy^2}\right)^2\right\}dy. \tag{9}$$

From equation (3) we have $\dfrac{dp}{dx} = \mu\left(\dfrac{\partial^2\bar{u}}{\partial y^2}\right) - \rho\dfrac{d}{dy}\left(\overline{u'v'}\right)$. Comparing this with (2a) we see that in I_1 we may take

$$\frac{du}{dy} = -\frac{3U_m}{b^2}y$$

since $\overline{u'v'}$ is small. Equation (8) reduces to

$$I_1 = \frac{6\pi U_m}{b^2l^2}\int_{-b}^{b}\left(\beta\frac{d\alpha}{dy} - \alpha\frac{d\beta}{dy}\right)y\,dy.$$

If turbulent motion is just possible then equation (8) may be written

$$\frac{3}{2}K_1 = \frac{lb^3\displaystyle\int_{-b}^{b}\left\{\left(\frac{2\pi}{l}\right)^4(\alpha^2+\beta^2) + 2\left(\frac{2\pi}{l}\right)^2\left[\left(\frac{d\alpha}{dy}\right)^2 + \left(\frac{d\beta}{dy}\right)^2\right] + \left(\frac{d^2\alpha}{dy^2}\right)^2 + \left(\frac{d^2\beta}{dy^2}\right)^2\right\}dy}{3\pi\displaystyle\int_{-b}^{b}\left(\beta\frac{d\alpha}{dy} - \alpha\frac{d\beta}{dy}\right)y\,dy} = b^3\Delta, \tag{9a}$$

where $K_1 = 2\rho bU_m/\mu$ and the function α, β, of y and the length l are to be determined so that Δ is a minimum.

Assuming
$$\begin{cases} \alpha = \alpha_0 + \displaystyle\sum_1^\infty (-1)^{n+1}\alpha_n\cos\frac{n\pi y}{b}, \\ \beta = \beta_0 + \displaystyle\sum_1^\infty (-1)^{n+1}\beta_n\cos\frac{n\pi y}{b}, \end{cases}$$

which satisfy the boundary conditions if $\alpha_0 - \sum_1^\infty \alpha_n = \beta_0 - \sum_1^\infty \beta_n = 0$. Taking $\alpha_i = \beta_i = 0$ for $i > 2$, substituting the values α and β in Δ we have

$$\Delta = \frac{\begin{aligned} &2b\left(\frac{2\pi}{l}\right)^4 (A^2 + P^2) + b\left(\frac{2\pi}{l}\right)^4 [(B^2 + Q^2) + (C^2 + R^2)] \\ &+ 2b\left(\frac{2\pi}{l}\right)^2 \left[\left(\frac{\pi}{b}\right)^2 (B^2 + Q^2) + \left(\frac{2\pi}{b}\right)^2 (C^2 + R^2)\right] \\ &+ b\left[\left(\frac{\pi}{b}\right)^4 (B^2 + Q^2) + \left(\frac{2\pi}{b}\right)^4 (C^2 + R^2\right] \end{aligned}}{\frac{2\pi}{l}\left[2b(AQ - BP) + 2b(AR - PC) + \frac{10b}{3}(CQ - BR)\right]},$$

where $\alpha_0 = A_1$, $\alpha_1 = B$, $\alpha_2 = C$, $\beta_0 = P$, $\beta_1 = Q$, $\beta_2 = \mathrm{R}$.
Putting $A = B + C$ and $P = Q + R$ we have

$$\Delta = \frac{\frac{\lambda_1}{2}(B^2 + Q^2) + \frac{\lambda_2}{2}(C^2 + R^2) + \mu(BC + QR)}{\nu(CQ - BR)}$$

where $\lambda_1 = 2\left[3\left(\frac{2\pi}{l}\right)^4 + 2\left(\frac{2\pi}{l}\right)^2\left(\frac{\pi}{b}\right)^2 + \left(\frac{\pi}{b}\right)^4\right]$, $\nu = \frac{20\pi}{3l}$,

$$\lambda_2 = 2\left[3\left(\frac{2\pi}{l}\right)^4 + 2\left(\frac{2\pi}{l}\right)^2\left(\frac{2\pi}{b}\right)^2 + \left(\frac{2\pi}{b}\right)^4\right], \quad \mu = 4\left(\frac{2\pi}{l}\right)^4. \tag{9b}$$

The conditions for a minimum value of Δ are

$$\begin{cases} B\lambda_1 + C\mu + R\nu\Delta = 0 \\ C\lambda_2 + B\mu - Q\nu\Delta = 0 \\ Q\lambda_1 + R\mu - C\nu\Delta = 0 \\ R\lambda_2 + Q\mu + B\nu\Delta = 0 \end{cases}$$

which have a solution provided the determinant

$$\begin{vmatrix} \lambda_1 & \mu & 0 & \Delta\nu \\ \mu & \lambda_2 & -\Delta\nu & 0 \\ 0 & -\Delta\nu & \lambda_1 & \mu \\ \Delta\nu & 0 & \mu & \lambda_2 \end{vmatrix} \equiv (\lambda_1\lambda_2 - \mu^2 + \Delta^2\nu^2)^2 = 0.$$

Substituting the value of λ_1, λ_2, μ, ν, from equations (9b) and minimizing this expression with respect to l we find $K = 131.4$. This is a value of the criterion when only three terms are used in the series α and β.

If we take $\alpha_i = \beta_i = 0$ for $i > 3$ we have

$$\alpha = A + B\cos\frac{\pi y}{b} - C\cos\frac{2\pi y}{d} + D\cos\frac{3\pi y}{b},$$

$$\beta = P + Q\cos\frac{\pi y}{b} - R\cos\frac{2\pi y}{b} + S\cos\frac{3\pi y}{b}.$$

Substituting these values in Δ we arrive at a determinant of the 6th order. This determinant is similar to that of the fourth order in the previous calculation and is also a perfect square. In fact, it can be easily proven that all succeeding determinants are perfect squares.

From the determinant of the 6th order we obtain a value for $K_1 = 120.9$ and from the 8th order $K_1 = 118.5$. Due to the mathematical computation involved this process becomes very laborious.

In the preceding work we expressed α and β by Fourier's Series which satisfied the boundary conditions exactly and determined conditions on the coefficients which were necessary for a minimum value of Δ and so found an approximate minimum for K_1. We can also determine by Calculus of Variations, differential equations which α and β must satisfy, if they furnish a minimum value of Δ, solve these equations and determine the arbitrary constants so that the boundary conditions are satisfied.

Beginning with the expression for Δ in (9) by the ordinary process of Calculus of Variations[4] we arrive at the differential equations,

$$2\left[\left(\frac{2\pi}{l}\right)^4\alpha - 2\left(\frac{2\pi}{l}\right)^2\frac{d^2\alpha}{dy^2} + \frac{d^4\alpha}{dy^4}\right] = -\Delta\left(\beta + 2y\frac{d\beta}{dy}\right)$$

$$2\left[\left(\frac{2\pi}{l}\right)^4\beta - 2\left(\frac{2\pi}{l}\right)^2\frac{d^2\beta}{dy^2} + \frac{d^4\beta}{dy^4}\right] = \Delta\left(\alpha + 2y\frac{d\alpha}{dy}\right).$$

Putting $\dfrac{\Delta}{2\left(\dfrac{2\pi}{l}\right)^4} = k_1$, and $y' = \dfrac{2\pi}{l}y$ and omitting $(')$, we have

$$\begin{aligned}\alpha - 2\frac{d^2\alpha}{dy^2} + \frac{d^4\alpha}{dy^4} &= -k_1\left(\beta + 2y\frac{d\beta}{dy}\right)\\ \beta - 2\frac{d^2\beta}{dy^2} + \frac{d^4\beta}{dy^4} &= k_1\left(\alpha + 2y\frac{d\alpha}{dy}\right).\end{aligned} \tag{10}$$

Multiplying the second equation in (10) by i adding it to the first and replacing ik by k, we have

$$(D^2 - 1)^2(\alpha + i\beta) = k(1 + 2yD)(\alpha + i\beta). \tag{11}$$

This equation has four linearly independent solutions of the form

$$\alpha + i\beta = P + kQ + k^2R + \ldots$$

where $P = \cosh y$, $\sinh y$, $y \cosh y$ or $y \sinh y$. The general solution contains four arbitrary complex constants. The boundary conditions which the arbitrary constants must satisfy separate them into two sets one corresponding to the solutions with $P = \cosh y$ and $y \sinh y$, the other to the solutions with $P = \sinh y$ and $y \cosh y$. The former are found to lead to a lower value of k. Assuming a power series in y for $\alpha + i\beta$ and substituting in (11) we find with $P = \cosh y$.

$$\alpha_1 + i\beta_1 = \left\{ \begin{array}{l} 1 + \frac{y^2}{2!} + (1+k)\frac{y^4}{4!} + (1+7k)\frac{y^6}{6!} + (1+22k+9k^2)\frac{y^8}{8!} + \\ (1+50k+109k^2)\frac{y^{10}}{10!} + (1+95k+583k^2+153k^3)\frac{y^{12}}{12!} + \\ (1+161k+2097k^2+2595k^3)\frac{y^{14}}{14!} + (1+252k+6006k^2+ \\ 19612k^3+3825k^4)\frac{y^{16}}{16!} + (1+372k+14574k^2+97732k^3+ \\ 82905k^4)\frac{y^{18}}{18!} + \ldots \end{array} \right.$$

and with $P = \dfrac{y \sinh y}{2}$

$$\alpha_2 + i\beta_2 = \left\{ \begin{array}{l} \frac{y^2}{2!} + \frac{2y^4}{4!} + (3+5k)\frac{y^6}{6!} + (4+28k)\frac{y^8}{8!} + (5+90k+ \\ 65k^2)\frac{y^{10}}{10!} + (6+220k+606k^2)\frac{y^{12}}{12!} + (7+455k+3037k^2+ \\ 1365k^3)\frac{y^{14}}{14!} + (8+840k+10968k^2+17880k^3)\frac{y^{16}}{16!} + \\ (9+1428k+32094k^2+122468k^3+39585k^4)\frac{y^{18}}{18!} + \ldots \end{array} \right.$$

where the coefficients are determined by the law

$$A_{n+4} - 2A_{n+2} + \{1 - (2n+1)k\}A_n = 0.$$

Taking the solution $\alpha + i\beta = (C_1 + iC_2)(\alpha_1 + iB_1) + (C_3 + iC_4)(\alpha_2 + i\beta_2)$,

then
$$\begin{aligned} \alpha &= C_1\alpha_1 - C_2\beta_1 + C_3\alpha_2 - C_4\beta_2 \\ \beta &= C_1\beta_1 + C_2\alpha_1 + C_3\beta_2 + C_4\alpha_2. \end{aligned}$$

The arbitrary constants C_1, C_2, C_3, C_4 must satisfy the conditions given by $\alpha = 0, \beta = 0, \dfrac{\partial\alpha}{\partial y} = \alpha' = 0, \dfrac{\partial\beta}{\partial y} = \beta' = 0$ when $y = \dfrac{2\pi b}{l}$. Eliminating C_1, C_2, C_3, C_4,

$$\begin{vmatrix} \alpha_1 & -\beta_1 & \alpha_2 & -\beta_2 \\ \beta_1 & \alpha_1 & \beta_2 & \alpha_2 \\ \alpha_1' & -\beta_1' & \alpha_2' & -\beta_2' \\ \beta_1' & \alpha_1' & \beta_2' & \alpha_2' \end{vmatrix} = 0.$$

The left-hand member is of the form $R^2 + S^2$ where

$$R + iS = \begin{vmatrix} \alpha_1 + i\beta_1 & \alpha_2 + i\beta_2 \\ \alpha_1' + i\beta_1' & \alpha_2' + i\beta_2' \end{vmatrix} \text{ and } R = S = 0. \qquad (12)$$

When the values of α_1, α_2, β_1, β_2, are substituted in (12) S is found to vanish identically and $R = 0$ gives the following equation:

$$0 = 1 + \frac{2y^2}{3!} + \frac{8y^4}{5!} + \frac{32y^6}{7!} + (128 + 32k^2)\frac{y^8}{9!} + (512 + 320k^2)\frac{y^{10}}{11!} + (2048 + 2816k^2)\frac{y^{12}}{13!} + (8192 + 21504k^2)\frac{y^{14}}{15!} + (32768 + 147456k^2 + 15360k^4)\frac{y^{16}}{17!} + (131072 + 933888k^2 + 276480k^4)\frac{y^{18}}{19!} + \ldots$$

where $y = \frac{2\pi b}{l}$. We wish to determine the value of y which gives the least value of $K_1 = \frac{4}{3}\left(\frac{2\pi}{l}\right)^4 b^3 k$. We therefore assume a series for $\frac{32k^2y^6}{9!}$ of the form

$$\frac{A}{y^2} + B + Cy^2 + Dy^4 + Ey^6.$$

Substituting and determining the coefficients A, B, C, D, E, we find

$$\frac{32k^2y^6}{9!} = -\frac{1.00556}{y^2} - 0.244949 - 0.04020465y^2 - 0.001426054y^4 + 0.00010464y^6.$$

which gives a minimum value of $K = 116.84$.

[1] Collected Papers, **2**, p. 524.
[2] Transactions, *Amer. Math. Soc.*, 1905.
[3] *Proc. Roy. Irish Soc.*, **27**, 1907.
[4] Wilson, *Advanced Calculus*, p. 400.

NORMAL DIVISION ALGEBRAS OF ORDER 2^{2m}

BY A. ADRIAN ALBERT

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated May 11, 1931

1. *Generalized Quaternion Algebras.*—Let F be any non-modular field and let

$$B = (1, i, j, ij), \quad ji = -ij, \quad i^2 = \alpha, \quad j^2 = \beta, \tag{1}$$
$$C = (1, x, y, xy), \quad yx = -xy, \quad x^2 = \gamma, \quad y^2 = \delta, \tag{2}$$

with $\alpha, \beta, \gamma, \delta$ in F be generalized quaternion algebras over F. The author have proved[1] that a necessary and sufficient condition that the direct product $A = B \times C$ be a division algebra is that the equation

$$\alpha\xi_1^2 + \beta\xi_2^2 - \alpha\beta\xi_3^2 = \gamma\xi_4^2 + \delta\xi_5^2 - \gamma\delta\xi_6^2 \tag{3}$$

be not solvable for $\xi_1, \ldots, \xi_6$ not all zero in F. As a simple application of the theory of Helmut Hasse[2] we now have

THEOREM 1. *Let $F = R(\theta)$, an algebraic field over the field R of all rational numbers. Then the direct product $A = B \times C$ of two generalized quaternion algebras over F is never a division algebra.*

As another application of Hasse's theorems we have

THEOREM 2. *Let ρ and σ be in $F = R(\theta)$ and let there exist quantities $\gamma_1, \ldots, \gamma_6$ in F such that $\gamma_2\gamma_4\gamma_6 \neq 0$ and*

$$(\gamma_1^2 - \gamma_2^2\rho)(\gamma_3^2 - \gamma_4^2\sigma) = \gamma_5^2 - \gamma_6^2\sigma\rho. \tag{4}$$

Then if $\alpha = \rho$, $\beta = \gamma_3^2 - \gamma_4^2\sigma$ in a generalized quaternion division algebra B of (1), algebra B contains a quantity t not in F but such that

$$t^2 = \Delta_1^2 + \Delta_2^2 \qquad (\Delta_1, \Delta_2 \text{ in } F). \tag{5}$$

As a further application the author has proved

THEOREM 3. *A necessary and sufficient condition that a normal division algebra A of order sixteen over an algebraic field $F = R(\theta)$ be a cyclic (Dickson) algebra is that A contain a quantity t not in F and satisfying an equation (5).*

Professor J. H. M. Wedderburn has proved[1] that a cyclic division algebra A of order sixteen over F defined by a cyclic quartic field $F(x)$ and a quantity γ in F is a division algebra if γ^r is not the norm $N(f)$ of any f in $F(x)$, for $r = 1, 2, 3$. The field $F(x)$ contains a quadratic field $F(u)$, $u^2 = \tau$ in F. We have now

THEOREM 4. *The Wedderburn norm condition for cyclic algebras of order sixteen is necessary as well as sufficient for $F = R(\theta)$, an algebraic field. It may be replaced by the simpler condition*

$$\gamma \neq \xi_1^2 - \xi_2^2 \tau$$

for any ξ_1 *and* ξ_2 *in* F.

But the author has also proved by an example

THEOREM 5. *Let* $F = R(\xi, \eta)$ *where* ξ *and* η *are indeterminates. There exist cyclic normal division algebras over* F *for which* $\gamma^2 = N(f)$, f in $F(x)$.

Hence the Wedderburn norm condition is, in general, not a necessary condition.

2. *Algebras of Order 16 over F.*—Consider a normal division algebra A or order sixteen over F. Such an algebra is known to contain a quartic field $F(x)$ which is the direct product of two quadratic fields $F(u)$ and $F(v)$ where $u^2 = \rho$, $v^2 = \sigma$, and ρ and σ are in F. Algebra A then[3] has a basis

$$j_r, \quad uj_r, \quad vj_r, \quad uvj_r \quad (r = 0, 1, 2, 3)$$

where

$$j_0 = 1, j_3 = j_1 j_2,$$

and

$$j_1^2 = \gamma_1 + \gamma_2 u, \quad j_2^2 = \gamma_3 + \gamma_4 v, \quad j_3^2 = \gamma_5 + \gamma_5 uv,$$

with $\gamma_1, \ldots, \gamma_6$ in F and such that

$$(\gamma_1^2 - \gamma_2^2 \rho)(\gamma_3^2 - \gamma_4^2 \sigma) = \gamma_5^2 - \gamma_6^2 \sigma \rho.$$

The author has now proved

THEOREM 5. *A necessary and sufficient condition that a normal division algebra* A *of order sixteen over any non-modular field* F *be a direct product of two generalized quaternion algebras over* F *is that there exist quantities* ξ_1, ξ_2 *in* F *such that*

$$\gamma_3^2 - \gamma_4^2 \sigma = \xi_1^2 - \xi_2^2 \rho. \tag{6}$$

Also the direct product

$$A \times A = P \times Q \tag{7}$$

where P *is a total matric algebra over* F *and*

$$Q = (1, i, j, ij) \quad ji = -ij, \quad i^2 = \rho, \quad j^2 = \gamma_3^2 - \gamma_4^2 \sigma \tag{8}$$

is a generalized quaternion algebra over F *which is a total matric algebra over* F *or a division algebra over* F *according as* A *is or is not expressible as a direct product of generalized quaternion algebras over* F.

For every normal division algebra A of order n^2 over F the *exponent* of A is defined to be the least integer ρ such that the direct power A^ρ is a total matric algebra. Theorem 5 then gives

THEOREM 6. *The exponent of a normal division algebra* A *of order sixteen over* F *is two or four according as* (6) *can or cannot be satisfied for* ξ_1 *and* ξ_2 *in* F.

Also Theorem 1 gives immediately

THEOREM 7. *The exponent of any normal division algebra of order sixteen over an algebraic field* $R(\theta)$ *is four.*

Finally, as an application of Theorems 7, 3, 2 the author has proved

THEOREM 8. *All normal division algebras of order sixteen over an algebraic field $R(\theta)$ are cyclic (Dickson) algebras.*

3. *Algebras of order 2^{2m}.*—Let A be a normal division algebra of order 2^{2m} over an algebraic field $R(\theta)$. The author has proved

THEOREM 9. *The direct product $A \times A$ is a total matric algebra if and only if $m = 1$.*

THEOREM 10. *The exponent of every A of order 2^{2m} is 2^m.*

THEOREM 11. *Let A and B be normal division algebras over an algebraic field $R(\theta)$ and let the orders of both A and B be powers of two. Then $A \times B$ is not a division algebra.*

Definition.—An equation $\phi(\xi) = 0$ is said to *belong to a normal division algebra A over F* if $\phi(\xi)$ has coefficients in F and is the minimum equation for F of a quantity x in A.

THEOREM 12. *Let A be a normal division algebra of order 2^{2m} over $F = R(\theta)$, an algebraic field over the field R of all rational numbers, and let $0 \leq t < m$, $s = 2^{m-t}$. Then*

$$A^s = M_s \times A_s$$

where M_s is a total matric algebra over F and A_s is a normal division algebra over F whose order is 2^{2t} and which is such that every equation of degree $2^r \leq 2^t$ which belong to A_s belongs to A and conversely.

THEOREM 13. *Every normal division algebra A of order 2^{2m} over $F = R(\theta)$, an algebraic field, contains a quantity x such that $F(x)$ has order 2^m and contains quantities*

$$x_m = x, \quad x_{m-1}, \ldots, x_1$$

with the property that if

$$K_0 = F, \quad K_i = F(x_i) \quad (i = 1, \ldots, m)$$

then K_i is a quadratic field over K_{i-1}.

THEOREM 14. *Every normal division algebra A of order 64 over an algebraic field $F = R(\theta)$ contains a sub-field which is the direct product of a cyclic quartic field and a field of order two over F, and hence A is an algebra of type R.*

The author has thus obtained a determination of all normal division algebras of order 64 over any algebraic field over R.

As another application the author proves

THEOREM 15. *A normal division algebra of order n^2 over an algebraic field $R(\theta)$ is a self-reciprocal algebra if and only if $n = 2$.*

Theorem 15 implies an extremely important result on pure Riemann matrices of the first kind. Such a Riemann matrix is one whose multiplication algebra is a normal division algebra over $R(\theta)$ such that the minimum equation of θ has all real roots. The only other case is where these roots are all imaginary, the Riemann matrices then being matrices of the second kind.[4] We now have

THEOREM 16. *The multiplication algebra of a pure Riemann matrix of the first kind is either a field $R(\theta)$ or a generalized quaternion algebra over such a field.*

[1] "On the Wedderburn Norm Condition for Cyclic Algebras," *Bull. Am. Math. Soc.*, **37**, pp. 301–312 (1931).

[2] "Quadratic Formen im algebraischen Zahlkorpern," *Jour. reine angewandete Math.*, **153**, pp. 113–130 (1923).

[3] *Cf.* the author's "New Results in the Theory of Normal Division Algebras," *Trans. Am. Math. Soc.*, **32**, pp. 171–195 (1930).

[4] See the author's paper, "The Structure of Pure Riemann Matrices with Non-Commutative Multiplication Algebras," *Rend. Circ. Mat. Pal.*, **55**, pp. 57–115 (1931), for the definitions and properties of Riemann matrices of the two kinds.

ELECTRIC CONDUCTIVITY AND OPTICAL ABSORPTION IN METALS, ONCE MORE

BY EDWIN H. HALL

JEFFERSON PHYSICAL LABORATORY, HARVARD UNIVERSITY

Read before the Academy, April 27, 1931

The question how many free electrons there are per cu. cm. of a metal is an important, if not indeed a vital, one in theories of electric conduction. Optical evidence has for twenty years been generally regarded as showing the number of free electrons to be about the same as the number of atoms. This belief, however, always assumes, so far as I have observed, that the free electrons are the only conduction electrons. But what happens to the optical argument if we once admit that electrons which are not "free," in the ordinary sense of the word, may have a part in conduction, by going directly from atoms to adjacent ions? I shall try in this paper to go somewhat more deeply into this question than I have done previously. I shall pay especial attention to the investigations, theoretical and experimental, of Meier[1] and of Hagen and Rubens.[2]

Meier uses the formulas of Voigt, but these are consistent with those of Drude, and I shall begin with Drude's general equation of motion for an electron, whether "free" or "bound," within a metal. It is equation (1) of Chapter 5 in his "Theory of Optics."

$$m\frac{\partial^2\xi}{\partial t^2} = eX - \frac{4\pi e^2}{\theta}\xi - re^2\frac{\partial\xi}{\partial t}. \qquad (1)$$

The term eX is the impressed electric force corresponding to the equation

$$X = X_0 \cos nt. \qquad (2)$$

The next term represents the elastic force tending to carry the electron back to its zero position. The final term represents the energy-absorbing, heat-producing, resistance encountered by the electron in its motion.

This resistance term is of a highly artificial character, for it represents the moving electron as encountering a continuous resistance proportional to the velocity. This conception of resistance, though fictitious, is convenient and permissible for the motion of electrons through steady fields, for it accords with the rate of production of joulean heat in such motions. But to use the same conception for the motion of electrons in the rapidly alternating fields of light-waves is a hazardous procedure which must affect with some measure of doubt all conclusions to which it leads.

Meier infers from his experiments and calculations that a very appreciable amount of energy is absorbed by the "bound" electrons, saying that in many cases they are of "prevailing influence." I do not undertake to criticize any of his conclusions regarding these electrons. But I raise the question whether the conduction electrons of which his experiments give evidence are necessarily, or probably, free electrons. I shall undertake in this paper to indicate the way in which the dual theory of conduction can deal with the matter before us.

Drude adapts equation (1) to the case of a free electron by making the second term, the elasticity term, in the second member zero. He thus gets

$$m \frac{\partial^2 \xi}{\partial t^2} = eX - re^2 \frac{\partial \xi}{\partial t}. \qquad (3)$$

Now the dual theory rejects this equation as applied to a free electron. In accordance with the usual conception of free-electron conduction it assumes that a free electron gains velocity, from an assisting electric field, during its free path, thus storing up increased kinetic energy to be turned into heat by collision at the end of the path. Moreover, this theory supposes the free electrons to be few, compared with the atoms, and their paths very long in duration compared with the period of a light-wave. Accordingly, many alternations of impulse will be applied to an electron during its path and these impulses will neutralize each other in their effect upon the velocity with which the electron impinges on an atom or an ion at the end of its path. The final term in equation (3) is meaningless for such a case. According to the dual theory the effect of the "free" electrons in optical absorption is negligible. Equation (3) should be interpreted as applying to the other class of conduction electrons, the "associated" electrons.

The elasticity term of equation (1) disappears for an associated electron, as it does for a free electron, but for a different reason. It disappears because the work done by elastic forces when an electron goes out from an atom to unite with an adjacent ion is zero, on the whole. As to the

final term in equation (3), the resistance term, it doubtless still gives an imperfect picture of the nature of resistance, but this picture is not so fatally false in this case as is that of a free electron. For the time of transit of an electron from an atom to a neighboring ion is probably of the same order of magnitude as the vibration-periods of "bound" electrons, which, according to the conclusions of Meier, are comparable with the periods of the light-waves used in his experiments. That is, the direction of the impressed electric field will not reverse many times, and may not reverse at all, during the period of an associated-electron "transit." Accordingly, the associated electrons can absorb energy from the field during their transits and generate heat at the expense of this energy when they make their impacts at the end of these transits.

Meier, studying the reflection of polarized light from various metals, found for each metal and for each wave-length of light used a value of n, the refractive index, and of κ, the coefficient of absorption. Using the nomenclature of Voigt, though with the fundamental ideas of Drude, he writes

$$2n^2\kappa = \frac{\rho_1 \nu_1'/\nu}{\nu^2 + \nu_1'^2} + \sum \frac{\rho\nu'\nu}{(\nu_0^2 - \nu^2)^2 + \nu'^2\nu^2}, \qquad (4)^3$$

and

$$n^2(1 - \kappa^2) = 1 - \frac{\rho_1}{\nu^2 + \nu_1'^2} + \sum \frac{\rho(\nu_0^2 - \nu^2)}{(\nu_0^2 - \nu^2)^2 + \nu'^2\nu^2}. \qquad (5)$$

In each of these equations the summation term has reference to the bound electrons. The term containing factors with the subscript 1 has to do with the conduction electrons.

Curves were plotted, from the observational data, showing $n^2\kappa$ and $n^2(1 - \kappa^2)$ as functions of the wave-length. Then came a laborious process of finding, by trial, values of ρ_1, ν_1', ρ, ν' and ν_0 which, when used in equations (4) and (5), would give values of $n^2\kappa$ and $n^2(1 - \kappa^2)$, according well with the observational curves already plotted for these quantities.

The quantity which especially interests us is ρ_1, defined as $4\pi N_1 e^2 \div m$, N_1 being the number of conduction electrons per cu. cm. of the metal. The process of evaluating N_1 numerically involved an estimate of $e \div m$ somewhat different from the one now accepted, but this difference is not important for our present purpose. Working by a method which he regards as unquestionably sound, Meier gets for each of the ten metals examined a value, p_1, for the number of conduction electrons per atom. This number varies from about 0.46, in copper, to about 2.22, in bismuth, the mean for all the metals examined being not far from 1.

Now in one aspect this average is in excellent agreement with the requirements of the dual theory of conduction; for it is natural to suppose that each atom can allow one of its electrons, and only one, to leave it

for transit to an adjacent ion. On the other hand, if Meier's results are to be interpreted as meaning that about one electron per atom, as a rule, is actually traveling at every instant in the direction of the current, this conception is not in accord with the picture of associated-electron conduction given by the dual theory. For this picture presents any one associated electron as quiescent, so far as conduction is concerned, for the greater part of the time, but sharing now and then, doubtless a vast number of times a second, in a sort of train movement through the metal. If there is any inconsistency between this view of the matter and Meier's conclusions regarding the number of conducting electrons, this inconsistency may perhaps reasonably be attributed to the highly artificial conception of resistance which is embodied in the equations used, a matter which has already been commented on in this paper.

But Meier estimated p by a second method, by taking, for each metal, the conductivity, per conduction electron, indicated by his investigation and finding how many electrons of this conductivity would be needed to account for the specific conductivity of the metal as measured with steady currents. He thus found values, which we will call p_2 values, that were notably larger in most cases than the p_1 values. He was surprised and puzzled by this fact but presently hit, somewhat doubtfully, on what is probably the correct explanation, for he remarks that "possibly a difference of damping constant—in conduction and oscillations plays a part here."

The present paper may be regarded as an endeavor to explain this difference which Meier, on the experimental evidence, surmised, but which he, with his conception of conductivity, had not anticipated and could not understand.

TABLE 1

METAL	p_1	p_2	$p_1 \div p_2$	λ_0 IN $\mu\mu$	MAX. $\frac{\lambda_0}{c}$	MIN. $\frac{\lambda_0}{c}$
Hg	2.13	2.19	0.97			
Bi	2.22	3.85	0.58	150		5×10^{-16}
Ni	0.66	3.85	0.17	305 to 150	1.02×10^{-15}	5×10^{-16}
Au	0.74	4.61	0.16	370 to 150	1.23×10^{-15}	5×10^{-16}
Ag	0.82	7.28	0.11	270, 245, 200	0.9×10^{-15}	6.7×10^{-16}
Co	1.08	10.8	0.10	525 to 150	1.75×10^{-15}	5×10^{-16}
Cu	0.46	5.9	0.078	500 to 200	1.67×10^{-15}	6.7×10^{-16}
Pt	1.33	27.5	0.048	600	2.00×10^{-15}	
Zn	0.89	18.9	0.047	750 to 150	2.50×10^{-15}	5.0×10^{-16}
Steel	1.32	32.6	0.040	750 to 200	2.50×10^{-15}	6.7×10^{-16}

Meier's material enables us to make table 1. The wave-lengths λ_0 are those corresponding to the frequency of the "chief characteristic vibrations" of the bound electrons. Accordingly, $\lambda_0 \div c$ is the vibration-period of a bound electron. The order of arrangement of the metals in this table is that given by Meier, who remarks that the less $p_1 \div p_2$ is for any metal, the greater, as a rule, is the value of maximum λ_0 for that

metal—that is, the greater the maximum vibration-period among the bound electrons of the metal. This observation seems to me important.

The fact that p_2 is greater than p_1 means that the conduction electrons are less effective for the alternating fields of the light-waves used than they are for steady fields. Can it be shown that such a difference should appear, with sufficiently short wave-lengths, if conduction is effected mainly by electron "transits" from an atom to an adjacent ion? An affirmative answer to this question seems to be indicated by the fact, noted by Meier, that the degree of failure, the amount of falling off, in the conductivity per electron in the alternating field of light-waves appears to have a pretty close relation to the natural vibration-period of the more loosely "bound" electrons. For it seems probable that the time occupied by an electron in "transit" is not very different from the vibration-period of these loosely bound electrons, since an outward swing of an outer electron would seem to be the natural beginning of a transit.

More than one authority has dealt with the ratio between steady-current conductivity and light-wave conductivity of metals. Thus Richardson, on page 432 of his "Electron Theory of Matter," gives a formula relating to this matter, crediting it to H. A. Wilson elaborating "a method originally given by Jeans;" but this formula is not in convenient shape for my use, as it does not explicitly contain the duration time of a "path." Thomson, on page 84 of his "Corpuscular Theory of Matter," says, "We can easily show that if k is the conductivity under steady forces, then when forces vary as sin nt the conductivity will be proportional to $k\,\frac{\sin^2 nT}{n^2T^2}$," where $2T$ is the interval between two collisions" [of a free electron with the atoms]. This formula I can readily adapt to my use by taking k to be the same as my "associated-electron conductivity," κ_a, and substituting duration of a "transit" for duration of a "path." Moreover, I let τ, instead of Thomson's $2T$, represent the duration of a transit. Thus I get for my present use $\kappa_a\,\frac{\sin^2 (1/2 n\tau)}{(1/2 n\tau)^2}$ as the adapted Thomson formula.

I must admit that I have never succeeded in my attempts to deduce Thomson's expression. Not being willing to use it blindly, I have invented a crude method of dealing with the problem in hand, a method which involves simplifying assumptions and can hardly be regarded as strictly accurate. This method has not given me a brief formula, and an exposition of it cannot well be attempted here, though it may be published in some future number of these PROCEEDINGS. Fortunately it gives numerical results agreeing, in all the cases where I have used it, to the third place with those derived from the adapted Thomson formula. In the following tabulation τ means the duration of a "transit," t_1 the com-

plete period of a light-wave; $\kappa_a' \div \kappa_a$ is the ratio of the reduced transit-conductivity to the normal, steady-current, transit conductivity.

TABLE 2

$t_1 \div \tau$	=	18	10	6	2	4/3	1
$\kappa_a' \div \kappa_a$	=	0.990	0.968	0.912	0.405	0.090	0

When τ is greater than t_1, the ratio $\kappa_a' \div \kappa_a$ will not necessarily be zero, but it will be small, at the most about 0.04, according to the Thomson formula.

Going back now to table 1 and remembering that, according to my conception, the $p_1 \div p_2$ of that table should be the same as the $\kappa_a' \div \kappa_a$ of table 2, we see that in order to account for the value $p_1 \div p_2 = 0.97$, the one given for mercury, the ratio $t_1 \div \tau$ should be about 10. Now the range of wave-length used with mercury was from $325\mu\mu$ to $630\mu\mu$, and accordingly the value of $t_1 \div \tau$, τ being constant and t_1 proportional to the wave-length, must have nearly doubled between the low and the high limits of t_1 for this range. At only one point could the value $t_1 \div \tau = 10$ have held. My theory would therefore lead me to expect a smaller value of $\kappa_a' \div \kappa_a$ near the short wave-length end than near the long wave-length end of the range used, though the difference might be too small for certain detection. Meier, however, made no allowance for any dependence of conductivity on wave-length. Moreover, he attributed none of the absorption in mercury to action of the bound electrons. In the case of this one metal, the conduction electrons, taken with their normal conductive power, were sufficient, he believed, to account for the observed facts, within the errors of observation and calculation. Nevertheless, an examination of the two curves which he gives for $n^2(1 - \kappa^2)$, as a function of λ, in the case of mercury (see his figure 40) shows that from wave-length $400\mu\mu$ to wave-length $325\mu\mu$ the two diverge in a way that may indicate a progressive failure of conductive power in the electrons.

As to the other metals dealt with, in all of which a reduced conductive power was indicated by the absorptive power attributed to the conduction electrons, it may be doubted whether the agreement between the experimental and the theoretical curves is close enough to prove the correctness of Meier's assumption that the absorptive power of these electrons is independent of the wave-length within the range of light employed. In many cases the two curves cross each other, sometimes more than once, a fact which at least suggests the possibility that closer agreement might have been obtained by assuming a variation of conductive absorptive power with variation of wave-length.

The smallest value of $p_1 \div p_2$ in table 1 is 0.04, for steel. This, according to table 2, corresponds to a value of $t_1 \div \tau$ lying between 4/3 and 1, but nearer the former. Let us call it x. The value of $t_1 \div \tau$ must

have been less than x for $\lambda = 225\mu\mu$ and greater than x for $\lambda = 600\mu\mu$. the limits of wave-length used with steel. It seems probable, then, that with a wave-length of 4μ, which is about 7 times $600\mu\mu$, the ratio $t_1 \div \tau$ would be near 10 and the ratio $\kappa_a' \div \kappa_a$ near 1. We may, then, conclude that all the metals listed in table 1 would have $\kappa_a' \div \kappa_a$ near 1 when $\lambda = 4\mu$.

Especial attention is directed to $\lambda = 4\mu$ because this was the short wave-length limit in the famous work of Hagen and Rubens[4] on absorption in metals. These investigators, using seven metals and several alloys, studied especially the quantity $(100 - R)$, in which 100 stands for the intensity of radiation falling at perpendicular incidence on a mirror of the metal and R represents the intensity of the radiation reflected. This quantity $(100 - R)$ is called the *penetrating radiation*. It has been shown by various writers[5] that Maxwell's original theory of radiation, when molecules and their motions are disregarded, leads to the general formula

$$(100 - R) = \frac{200}{\sqrt{\Lambda t_1}} \qquad (6)$$

where Λ is the electric conductivity of the metal in absolute electrostatic measure, and t_1 is the full period of the radiation wave in seconds. Letting

TABLE 3

	$\lambda = 4\mu$	$\lambda = 8\mu$	$\lambda = 12\mu$
Mean value of C for 4, 8, 12μ,	$C_4 = 19.4$	$C_8 = 13.0$	$C_{12} = 11.0$
Mean deviation from mean value of C	$\delta_4 = 21.0\%$	$\delta_8 = 14.5\%$	$\delta_{12} = 9.6\%$
Theoretical value of C' from eq. (7)	$C_4' = 18.25$	$C_8' = 12.9$	$C_{12}' = 10.54$

κ stand for the conductivity in the units employed by Jäger and Dissel-horst,[6] and putting λ for the wave-length in μ, Hagen and Rubens get from equation (6)

$$(100 - R)\sqrt{\kappa} = \frac{36.5}{\sqrt{\lambda}} = C, \qquad (7)$$

C being a quantity which, for a given value of λ, should be the same for all metals.

Working with wave-lengths 4μ, 8μ and 12μ, H. and R. obtained with the metals (except bismuth) which they studied, and five alloys, results agreeing fairly well with equation (7). They made the test in the following way: Multiplying the observed $(100 - R)$ of each material by $\kappa^{1/2}$, κ being the steady-current conductivity of the metal, and taking the mean of all the products, except that for bismuth, for each of the three wave-lengths, they called this mean C. From the second member, $36.5 \div \lambda^{1/2}$,

of equation (7) they got for each λ a value which they called C'. Table 3 is taken, with slight modifications of form, from their table on page 884.

I have certain observations to make relative to the numbers in this table. For the 4μ case $C \div C' = 1.063$. This ratio could have been made 1.000 by using in finding C a value k about 12 per cent less than the normal steady-current value. A natural interpretation of this fact would be that, in spite of the inference above drawn from the behavior of the metals studied by Meier, 4μ is a somewhat too short wave-length to show the full absorptive power to be expected from the conduction electrons. Confirmation of this view might seem to be found in the 8μ case, where $C \div C'$ is 1.008, and could have been made 1.000 by using in the calculation of C a value of κ only 1.6 per cent below the steady current value. But when we go on to the 12μ case we have $C \div C' = 1.044$, and to make it 1.000 a value of k nearly 9 per cent below the normal value would have been required.

It appears, then, that we do not find an unquestionable approach to the condition $C \div C' = 1.000$ as we increase the wave-length. The mean value of $C \div C'$, for the three wave-lengths, is 1.038, corresponding to a value of 1.077 for $\kappa \div \kappa'$, and there is a possibility that some such ratio as this would continue to much longer wave-lengths. If this should prove to be the case, the discrepancy might be explained in either of two ways. First, one might suppose that imperfection of the mirror surfaces of the metals made R too small and $(100 - R)$ too large. Second, some consideration might be given to the fact that the dual theory of conduction indicates a discrepancy in the direction noted and of something like the amount noted.

For it is to be remembered that, according to this theory, the κ that is effective in light absorption is not, even with long wave-lengths, the full steady-current value of the conductivity. It is merely what I call κ_a, and so it is κ_a, not the full κ, that should be used in calculating C. Excluding bismuth, because H. and R. exclude it in taking their mean values, I, from my own study of the electrical and thermal properties of 17 metals, including 2 alloys, find 1.096 as the mean value of $\kappa \div \kappa_a$ at 0° C. According to this, use of κ, instead of κ_a, as a factor in the value of C would make the ratio $C \div C'$ between 4 and 5 per cent greater than it should be. This is a little more than enough to account for the value, 1.038, found by H. and R. for the mean value of this ratio.

The metals and alloys of my list are not just the same as those used by H. and R., and I wish to guard against attaching too much importance to the numerical relations here noted. The main point which I hope to make in this connection is that the experimental evidence is not unfavorable to my conception of conduction.

A like remark is justified when we consider the peculiar behavior of

bismuth. The mean value of $C \div C'$ for bismuth, according to the work of H. and R., is excpetionally large, being 1.37. The value of $\kappa \div \kappa_a$ in bismuth is, according to my investigation, exceptionally large, being 1.51 at 0° C. Use of κ instead of κ_a would account for 1.23 as the value of $C \div C'$ in bismuth, to be compared with the mean value, 1.038, found by H. and R. for metals and alloys in general. H. Murmann (*Zeits. Z. Physik*, **54** (1929)) using wave-lengths ranging from 25μ to 110μ measured the amount of radiation actually transmitted by thin films of various metals, including bismuth. But his results have little bearing on the question here at issue, for neither the thickness nor the conductivity of a film was known, but only the product of these two quantities.

Summary

1. The resistance term in Drude's fundamental equation of motion of an electron within a metal is highly artificial, for light-wave fields, and any conclusions to which it leads are affected by some measure of doubt.

2. The form which this equation takes when it is applied to a conduction electron is quite as appropriate for an electron that passes directly from an atom to an adjacent ion as it is for a "free" electron.

3. The conduction electrons which Willi Meier (in 1909) found to be about as numerous as the atoms may reasonably be taken to be the "associated" electrons, the electrons that may execute "transits" from an atom to an adjacent ion.

4. The formula $k\frac{\sin^2 nT}{n^2T^2}$, given by J. J. Thomson on page 84 of his "Corpuscular Theory of Matter," when applied to "transit" conductivity gives the same results as those obtained in a different way by the author, indicating that when the ratio *wave-period to transit-period* drops from 10 to 1, the absorptive power of the transit electrons drops from 97% of its maximum value to zero.

5. Meier's results, obtained with wave-lengths ranging from $250\mu\mu$ to $670\mu\mu$, seem to indicate that for wave-lengths of 4μ or greater the transit electrons should have almost their full, steady-current, conductive power and a corresponding absorptive power.

6. Hagen and Rubens, using wave-lengths 4μ, 8μ and 12μ, got for metals and alloys in general results which seem to indicate that the absorption-conductivity is somewhat less than the steady-current conductivity at all three wave-lengths, the difference being about 12% at 4μ, about 1.6% at 8μ and about 9% at 12μ, with no clear evidence that it would disappear at greater wave-lengths. In the case of bismuth the difference is exceptionally large.

7. This difference, which is hardly to be accounted for by the ordinary theory of the relation between conduction and absorption, may be due to some imperfection of the experimental method used by H. and R. On the other hand, it is worth noting that the dual theory of conduction requires a difference of the kind and of the order of magnitude here observed. For this theory makes the paths of the few free electrons last so long that the rapidly alternating fields of even the longest light-waves would have no net effect on them. Therefore it confines the conduction-absorption power of metals to the action of the "associated," or "transit," electrons, and for metals in general the conductivity due to these electrons is about 8 or 10% less than the total conductivity. For bismuth κ_a is exceptionally small, about 67% of the total κ.

[1] Meier, *Inaugural Dissertation*, Barth, Leipsic, 1909.

[2] Hagen and Rubens, *Ann. d. Physik*, **11**, 873–901 (1903).

[3] Meier's paper gives $\rho\nu' \div \nu$ in the numerator within the summation sign, but I think it should be $\rho\nu'\nu$.

[4] Hagen and Rubens, *Ann. d. Physik*, **11**, 873–901 (1903).

[5] Loc. cit., p. 886.

[6] Their κ is the reciprocal of the resistance, in ohms, of a piece of the metal 1 m. long and 1 sq. m. in area of cross-section.

THE SPECTRAL ERYTHEMIC REACTION OF THE HUMAN SKIN TO ULTRA-VIOLET RADIATION

BY W. W. COBLENTZ, R. STAIR AND J. M. HOGUE

NATIONAL BUREAU OF STANDARDS, WASHINGTON, D. C.

Read before the Academy April 27, 1931

Introduction.—At present the physician has neither a unit of dosage nor a meter for accurately measuring the amount of ultra-violet radiation used for healing purposes.

In the absence of an inanimate dosage-meter the patient is used as an indicator, and the dosage is estimated by the erythema produced on the inner forearm. This method is in common use as a guide in ultra-violet radiation therapy. Furthermore, in view of the wide variation in erythemic susceptibility of pigmented and unpigmented skin (brunette and blonde), and in view of the fact that irradiation cannot be continued safely beyond skin tolerance, it is highly probable that any unit of dosage or any inanimate dosage-meter, that may be adopted, will have to take this physiological effect into consideration.

Hence, in connection with the question of the unit of dosage, and particularly in connection with methods of standardizing the dosage, an

important problem is the determination of the spectral erythemic reaction of the untanned human skin, and the energy required to produce a mild erythema.

Previous determinations of the erythemic response of the human skin when exposed to ultra-violet radiation of different wave-lengths were made by Hausser and Vahle[1] and more recently by Luckiesh[2] and his collaborators. In their determination of the energy required to produce a minimum perceptible erythema, Hausser and Vahle exposed the arm to single emission lines, which passed through a slit (3 cm. long) in a cardboard attached to the arm. Based upon their preliminary observations of the approximate magnitude of the erythemal dose for each wave-length, by making the exposure through the whole length of the slit, then through $^2/_3$ and finally through $^1/_3$ of the length, for different time intervals (3 exposures with each line), the exposure-time for producing a distinctly perceptible erythema was determined.

Instead of using isolated spectral lines, Luckiesh[2] and his collaborators exposed small areas of the subject's back to the total radiation from a quartz mercury arc lamp, shining through filters that shut off successively increasing wave-lengths of ultra-violet radiation, and noting the dosage necessary to produce a minimum perceptible erythema. For this purpose the time was varied so that the shortest exposure gave no visible results, and the longest exposure produced a definite erythema, as observed after a lapse of approximately 24 hours.

Experimental Procedure.—In the present investigation isolated spectral lines from a large quartz monochromator[3] were used as stimuli, and the time of exposure to produce a minimum perceptible erythema (one that lasted less than 24 hours) was determined. As found by others, the difficulty lies in defining the minimum perceptible erythema, especially the highly transitory erythema produced by wave-lengths less than 270 $m\mu$ (millimicrons) which, for a weak dosage, increases to a maximum redness and then disappears in the course of 2 to 5 hours.

The source of ultra-violet radiation used in the present work was a quartz mercury vapor lamp, operated on a constant voltage, under which condition the intensities of the spectral lines, measured at different times with a linear thermopile[5] calibrated in absolute value,[6] were constant to 5%.

The exit slit (of metal 0.1 mm. in thickness, with beveled edges turned away from the incident light[4]) was mounted upon a rigid support which, after focusing upon the desired spectral line, was securely attached to the spectrometer arm.

To facilitate setting on a spectral line the slit was given a thin coat of turpentine upon which was dusted a thin deposit of anthracene, which was then rubbed smooth. Starting with the slit closed this method of applying

the fluorescent material produces a sharp knife-edge which, on opening the slit (0.3 mm. wide, 10 mm. long), permits an accurate isolation of closely adjoining emission lines.

The arm to be exposed to ultra-violet radiation rested upon a movable support directly back of, and within less than 0.3 mm. from the front surface of, the exit slit. The difference in focus of the radiation incident upon the exit slit, and incident upon the skin directly back of the slit,

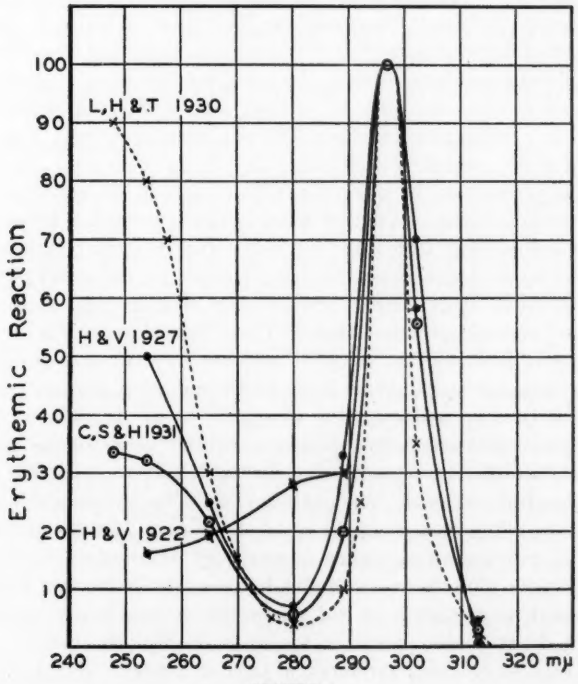


FIGURE 1

Spectral erythemic response of the human skin; H & V, Hausser and Vahle; L, H & T, Luckiesh, Holladay and Taylor; C, S & H, Coblentz, Stair and Hogue.

was therefore negligible. By moving the arm-support laterally, different parts of the arm (usually at intervals of 4 to 6 mm.) could be exposed to radiation.

In practice the exit slit was adjusted to intercept a given emission line (e.g., 297 mμ) and a series of irradiation exposures was made, ranging from 1.5, 1.75, 2.0, 2.25, 2.50 and 3.0 minutes. From such a series of exposures the time for producing a minimum perceptible erythema (1.75,

2.0 and 2.5 min., respectively, for the three writers) was easily established.

When making irradiation tests in the region of 254 mμ, where the observations of the erythemic reaction are in disagreement with previous experimenters, check exposures were made on the 297 mμ line alongside with the exposures to the emission line (e.g., 248 mμ) under observation. Since the minimum perceptible erythema for the 297 mμ line was produced in the usual time (1.75 min.; 2 min. produced an erythema lasting 2 days for W. W. C.) it is assumed that this disagreement is real.

The Spectral Erythemic Reaction.—As illustrated in figure 1, the wave-length range of the erythemogenic rays begins at about 315 mμ, and extends to an undetermined wave-length shorter than 240 mμ. The erythemic response curve rises abruptly to a maximum at 297 mμ, descends less abruptly to a minimum at about 280 mμ, then rises to a less intense maximum in the region of 250 mμ.

The disagreement in the spectral response curve in the region of 250 mμ is partly owing to the method of defining the minimum perceptible erythema, which is very transitory for short wave-lengths. However, this appears to be of minor importance in view of the increased practice of covering artificial light sources with a glass that intercepts ultra-violet radiation of wave-lengths less than 280 mμ, thus reducing the erythemogenic range of wave-lengths to agree more nearly with that of the sun.

The erythemogenic power of ultra-violet radiation depends upon the intensity of the rays and upon the susceptibility of the skin to different wave-lengths. But unlike the spectral sensibility curve of the eye which is different for different persons, from the limited data available it appears that the spectral erythemic response curve of the human skin is practically the same for different persons, in spite of the fact that the total energy required to produce an erythema is markedly different.

One difficulty in devising a suitable dosage meter is the fact that a long over-exposure to radiation at 254 mμ produces only a superficial burn, whereas a slight over-exposure to the more deeply penetrating radiation of wave-lengths 313 mμ produces a painful blister. For example, in one case it was found that an exposure to the wave-length 313 mμ for 14 min. produced a barely perceptible erythema (estimated exposure 15 min.); 16 min. was somewhat longer than required for a minimum perceptible erythema; and an exposure of 18 minutes produced a painful blister.

Reciprocity Law.—According to Luckiesh[7] the reciprocity law holds for biological effectiveness over a wide range of exposures.

In the present experiments the intensity of the emission line at 297 mμ was reduced by $^1/_2$, and by $^1/_4$, by means of glass filters, and the corresponding time of exposure was doubled and quadrupled. Within the experimental errors involved in making the test, the erythema produced

by using $^1/_4$ of the unit intensity and exposing for 4 times the unit time, was not appreciably different from that produced by exposing the subject to unit intensity (430 microwatts per cm.^{2}) for the unit time (2 min. for R.S.) required to produce a minimum perceptible erythema.

On the basis of the limited number of subjects tested in this work, the energy required to produce a minimum perceptible erythema upon 1 square centimeter of untanned skin, using monochromatic radiation of wave-length 297 mμ (the wave-length of maximum of erythemic susceptibility) is of the order of 500,000 ergs. The actual values for the three observers are, respectively, 452, 516 and 645 kiloergs per cm.^{2}

When one considers the ever-increasing number of ultra-violet radiators sold for healing purposes, in some of which about 97% of the total erythemogenic radiation, of wave-lengths 200 to 315 mμ is contained in the highly germicidal but non-penetrating rays at 254 mμ (the resonance line in a so-called "cold" quartz mercury lamp), while other lamps contain practically no ultra-violet radiation of wave-lengths shorter than 310 mμ but emit longer wave-lengths that are more penetrating and more potent in causing burns, the urgent need of a means of standardization of the dosage seems apparent.

[1] Hausser, K. W., and Vahle, W., *Strahlentherapie, Berlin*, **13**, 41–71 (1921), and **28**, 25–44 (1928).

[2] Luckiesh, M., Holladay, L. L., and Taylor, A. H., *Jour. Opt. Soc. Amer.*, **20**, 423–432 (1930).

[3] Coblentz, W. W., and Kahler, H., *Bur. Stds. Sci. Papers*, **16**, 233–248 (1920).

[4] Coblentz, W. W., *Jour. Franklin Inst.*, **175**, 151–153 (1913).

[5] Coblentz, W. W., *Bur. Stds. Bull.*, **11**, 131–187 (1914).

[6] Coblentz, W. W., *Ibid.*, **11**, 87–100 (1914).

[7] Luckiesh, M., "Artificial Sunlight," p. 77 (1930); *Trans. Illum. Eng. Soc.*, **26**, June-July (1931).